Developing an Online Predictor to Predict Product Sulfur Concentration for HDS unit

by

©Prafull Sharma

A Thesis submitted to the

School of Graduate Studies

in partial fulfillment of the requirements for the degree of

Master of Engineering

Faculty of Engineering and Applied Science

Memorial University of Newfoundland

May 2018

St.John's Newfoundland and Labrador

Abstract

Hydrodesulfurization (HDS) is an important process in refining industries. Advanced control system (e.g. model predictive controller) requires on-line measurement of the product sulfur at the reactor outlet. However, most HDS processes do not have a sulfur analyzer at the reactor outlet. In order to predict product sulfur concentration usually a data based sulfur predictor is developed. Performance of data based predictor is usually poor since some of the input parameters (e.g. feed sulfur concentration) are unknown. The objective of this thesis is to overcome these limitations of data based predictors and develop an online product sulfur predictor for HDS unit. In this thesis, a hybrid model is proposed, developed and validated (using industrial data), which could predict product sulfur concentration for online HDS system. The proposed hybrid structure is a combination of a reaction kinetics based HDS reactor model and an empirical model based on support vector regression (SVR). The mechanistic model runs in off-line mode to estimate the feed sulfur concentration while the data based model uses the estimated feed sulfur concentration and other process variables to predict the product sulfur concentration. The predicted sulfur concentration can be compared with the lab measurements or sulfur analyzer located further downstream of the process at the tankage. In case there is a large discrepancy, the predictor goes to a calibration mode and uses the mechanistic model to re-estimate the feed sulfur concentration. The detailed logic for the online prediction is also developed. Finally a Matlab based Graphical User Interface (GUI) has been developed for the hybrid sulfur predictor for easy implementation to any HDS process.

Tags: Hydrodesulfurization, Hybrid model, SVR, Online prediction, Design of Experiments, Graphical User Interface (GUI).

To my Nanaji, Nani, Mum and Dad

Acknowledgements

First of all, I would like to thank my supervisor Dr. Syed Imtiaz and co-supervisor Dr. Salim Ahmed for their valuable guidance and generous support. Dr. Imtiaz gave me the freedom to work at my pace and helped me keep going when the going got tough. I would like to thank my professors for teaching me various aspects of Design of Experiments and Process engineering. I thank my group mates Anirudh Nadan, Amin and Eugenio Turco for all his help and assistance. I thank Moya Crocker, Colleen Mahoney, and Nicole Parisi for creating a friendly and enabling atmosphere at the University.

I express my greatest gratitude to my parents & grandparents for being a constant source of inspiration, love and affection. I dedicate this thesis to them.

Finally, I would like to thank the School of Graduate Studies (SGS) for their generous financial support.

Table of Contents

List of Figures

List of Tables

Nomenclature

a = Dimensionless number of Glaso's correlation

a_j = specific surface area of the interface j (cm^2)

A = Aromatic Compound

API = API gravity

B = Saturated hydrocarbon

c_j^p cpj = Specific heat capacity of J phase ($J*g^{-1}*k^{-1}$)

c_j^i = Molar concentration of i in j phase ($mol*cm^{-3}$)

d_p= particle diameter (cm)

d_{15} = Liquid density at 15^0C ($g*cm^{-3}$)

$d_{15.6}$ = specific gravity at 15.6 ^{0}C

D_j^i= Molecular diffusivity of i in j phase (cm^2*sec^{-1})

E_a = Activation energy ($J*mol^{-1}$)

G_l= Liquid Superficial mass velocity ($g*cm^{-2}*sec^{-1}$)

H_i = Henry law constant for compound i($MPa*cm^3*mol^{-1}$)

k_L = thermal conductivity of liquid phase ($J*s^{-1}*cm^{-1}*K^{-1}$)

k_r = Reverse HDA rate constant (sec^{-1})

k_j^i = Mass transfer coefficient of compound I at the interface j ($cm*sec^{-1}$)

k_{H_2S}= Adsorption equilibrium constant for H2S (cm^3*mol^{-1})

L_b = length of catalyst bed (cm)

p_j^i = partial pressure of compound I in the j phase (MPa)

P = Reactor total pressure (psia)

r_j = Reaction rate j ($mol\ cm^{-3}\ sec^{-1}$ fpr j = HDS $mol*g^{-3}*s^{-1}$)

R = universal gas constant ($J*mol^{-1}*K^{-1}$)

t = time (sec)

T = temperature (K)

T_{MeAB} = Mean Average Boiling Point (^0K)

u_j = Superficial velocity of j phase ($cm*sec^{-1}$)

z = Axial coordinate (cm)

Greek symbols

ΔH_{R_j} = Heat of reaction j ($J*mol^{-1}$)

$\Delta\rho_T$ = temperature correction of liquid density ($kg*m^{-3}$)

$\Delta\rho_p$ = Pressure dependence of liquid density ($kg*m^{-3}$)

ϵ = Bed void fraction

ϵ_g, ϵ_l = Hold up for gas and liquid phase

ϵ_p = particle porosity

η_j = catalyst effectiveness factor for reaction j

λ_i = solubility coefficient of the compound i ($Nl*kg^{-1}*MPa^{-1}$)

ρ_B = catalyst bulk density ($g*cm^{-3}$)

ρ_j = density at Process condition of j Phase ($g*cm^{-3}$)

ρ_0 =liquid density at standard condition (@15.6^0C 101.3kPa) ($g*cm^{-3}$)

ρ_{20} = liquid density at 20^0C ($g*cm^{-3}$)

μ_l = Absolute viscosity of the liquid (MPa*sec)

υ_c = critical specific volume of the gaseous compound

υ_i = molar volume of solute i at its normal boiling temperature (cm^3*mol^{-1})

υ_L= molar volume of solvent liquid at its boiling temperature (cm^3*mol^{-1})

υ_N = molar volume at standard conditions

υ_{cm} = Critical specific volume (m^3*kg^{-1})

τ = fractional volume of catalyst bed diluted by inert particle

Chapter 1

1 Introduction

1.1 Importance of Desulfurization Process

Petroleum consists of a significant part of our lives and will do so for upcoming decades. The fuel that is produced from petroleum, supplies more than half of the world's total supply of energy. Gasoline, kerosene and diesel oil provide fuel for automobiles, aircrafts and ships. Fuel oil, heavy oil and natural gas are employed for heating homes and commercial buildings in addition to generating electricity. The declining reserves of oil in the world have resulted in decline in quality of petroleum, which has caused the nature of crude oil refining to change considerably (Speight et al., 2015).

The fuels extracted these days have high traces of sulfur, nitrogen, aromatic and various other compounds. The presence of sulfur and other compounds in crude oil has a significant impact upon quality of oil products. On combustion of such fuels, sulfur compounds releases into atmosphere as harmful sulfur oxides (SO_x) and sulfate particulates. These SO_x compounds are hygroscopic and in atmosphere, it reacts with humidity to forms sulfuric and sulfurous aerosol acid (acid rain). The exposure to sulfates and to acids derived from SO_x compounds can cause corneal haze, breathing difficulty, airways inflammation, eye irritation, psychic alterations, pulmonary edema, heart failure and circulatory collapse. Sulfur compounds can also cause asthma, chronic bronchitis,

morbidity and mortality increase in old people and infants. These compounds are noxious venom for vegetation when compared to human health. For example, the level of 0.3 SO_x µg per cubic meter of air implies a potential risk for human health, but for plants 0.2 SO_x µg per cubic meter of air is lethal. Therefore, sulfur oxides (SO_x) and sulfuric acid (H_2SO_4) causes damage and destruction of vegetation, soil deterioration, corrosion of construction material and erosion of watercourses (De la Paz-Zavala et al., 2013; Khalfhallah et al., 2009; Mederos et al., 2007). Usage of high sulfur concentration fuel can also cause corrosion of automobile engine, affecting its performance and increasing it maintenance. Studies have shown that amount of SO_x emissions is directly proportional to the amount of sulfur contained in the fuel (Kim et al., 2013). The SO_x oxides, particulates and other harmful pollutants present in atmosphere are main reasons for targeting sulfur content in fuel and decreasing it to lower levels in many countries worldwide. It has been proved that low level sulfur fuels and use of appropriate particulate filters can reduce particulate matter by as much as 90% (Marafi et al., 2007). Therefore, desulfurization of fuel oils is an important process in oil refinery.

In order to limit the concentration of SO_x, various countries have now stringent specification on petroleum products including diesel. The Environmental Protection Agency (EPA) has changed the allowable to limit sulfur in diesel fuels from 500 ppm (low sulfur diesel (LSD)) to 15ppm (ultra-low sulfur diesel (ULSD)) in 2010 (Ocic et al., 2010). Similar regulations have been proposed by European Commission (Jones et al., 2006), while the same path is followed by other legislations, introducing, limitations on

sulfur concentration in fuels. In EU, by the year 2006, the sulfur content was reduced to 50 ppm. By January of 2011 all on-road diesel in EU had a specification of maximum 10 ppm sulfur. In the EU, Germany was the first country to adopt the 10 ppm sulfur limit. Other EU countries and Japan introduced diesel fuel with 10 ppm soon after Germany. (EU fuel Regulation, 2016, Japan Fuel Regulation, 2016). Thus low sulfur concentration fuel has a demand in current market, making desulfurization as an important process.

1.2 Brief description of Different method for desulfurization

Desulfurization processes can broadly be classified in two groups, 'hydrotreating based' and 'non-hydrotreating based', desulfurization depending on the role of hydrogen in removing sulfur (Figure. 1.1) (Babich et al.,2003). In hydrotreating based process hydrogen is used to decompose organosulfur compounds and eliminate sulfur while non-hydrotreating based process uses other means to remove sulfur.

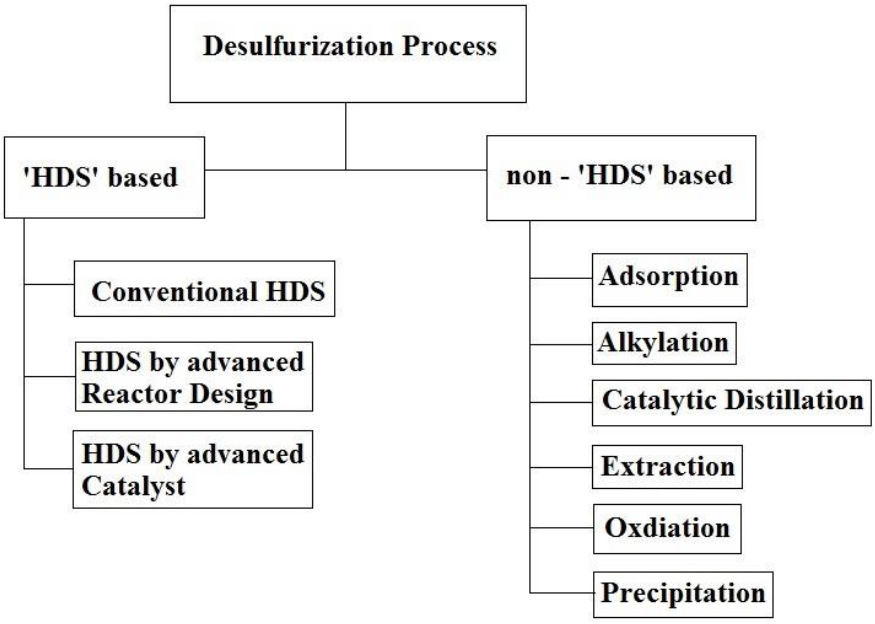

Figure 1.1: Desulfurization technology classified by nature of a key process to remove Sulfur (Babich et al., 2003)

1.3 Hydrodesulfurization process

Among all the various methods of desulfurization, hydrodesulfurization process (conventional HDS as per Figure 1.1) is commonly used to produce ultra-low concentration sulfur fuel. In hydrodesulfurization process, hydrogen gas reacts with liquid fuel with high sulfur concentration on a Nickel – Molybdenum catalyst bed to produce low sulfur concentration fuel with hydrogen sulfide gas. The overall reaction is stated as follows

$$(\mathrm{R-S}) + \frac{3}{2}(\mathrm{H_2}) \rightarrow (\mathrm{R-H}) + (\mathrm{H_2S})$$

As all the three phases are present (i.e. $\mathrm{H_2}$ in gaseous phase, fuel with high concentration of sulfur reacting, in a liquid phase on a Ni-Mo catalyst bed, solid phase) during the HDS

reaction, trickle bed reactor is used for HDS reaction (Farahani et al., 2011). In industrial scenario, as HDS unit is generally at the end of the process, effluents from various different units, having different sulfur concentration are passed through HDS unit to decrease sulfur concentration in fuel. This causes feed sulfur concentration values to vary in HDS unit. This results in fluctuation in product sulfur concentration accordingly, which may or may not result into a fuel with desired sulfur concentration. Thus, a predictor task is to predict the product sulfur concentration for the given feed conditions to HDS unit. Based on its estimation of product sulfur concentration controller will take necessary actions in order to ensure that the product sulfur concentration produced has always have acceptable sulfur levels.

1.3.1 Historical development of HDS process models

Various sulfur predictor have been designed which are able to predict the product sulfur concentration for a HDS unit. Initially, steady state models (mathematical models) were developed which described and captured HDS unit (Babich et al., 2003). These models were one dimensional heterogeneous model, using kinetic rate equations to describe HDS. These proposed models were able to predict product sulfur concentration in steady state. However, mathematical models developed in recent times, were able to capture dynamic behavior of HDS unit while predicting product sulfur concentration (Mederos et al., 2009; Deng et al., 2010; Cicili et al., 2009). Mathematical models can't predict for online simulations due to their shortcomings, therefore data driven model were used to develop model for HDS unit. (Lukec et al., 2008). Data driven models has its advantages

over mathematical model and these models are able to capture dynamic behavior of HDS unit in order to predict product sulfur concentration. (Kadlec et al., 2009; Lukec et al., 2008). When it comes to online prediction even data driven model has its limitations, as these models are only limited to the ranges they are trained for. Thus, the data used to trained these data driven model checked for any data outliners and noise, in order to minimize the modelling error. Thus in recent times, hybrid models are being used to predict the product sulfur concentration. These models are able to extrapolate and interpolate more accurately than data driven or mathematical models while estimating product sulfur concentration. (Bhutani et al., 2006).

1.3.2 Problem encounter / formulation

There are various sulfur predictors developed in past (Von Stosch et al., 2014), which can capture the dynamic behavior and can predict product sulfur concentration for HDS unit. These proposed sulfur models predicts product sulfur concentration satisfactorily for offline prediction. However, for an industrial perspective, it is necessary to predict sulfur concentration for HDS unit on an online system, which is a hassle because

- The feed sulfur concentration keeps on fluctuating, as the effluents from various different units having different sulfur concentrations are passed through HDS unit. In absence of sulfur analyzer at the beginning of HDS unit, no data is being recorded regarding feed sulfur concentration. Therefore, for any online predictor, predicting product sulfur concentration in absence of feed sulfur concentration data is quite difficult, especially when feed sulfur concentration keeps on varying.

- Unlike the input stream of HDS reactor, there is a sulfur analyzer installed at the output stream for the HDS unit. It records the product sulfur concentration and sends it back to the online predictor with a time lag. This causes hindrance in predicting product sulfur concentration effectively in a continuous manner.

In this thesis, our objective is to develop a sulfur predictor, when implemented on an online system for prediction will not only be able to handle above mentioned problems but also can predict product sulfur concentration effectively.

1.4 Objective of this thesis

Based on the problem stated above, in this thesis, a sulfur predictor is developed which estimates product sulfur concentration for a HDS unit when implemented on an online system. As numerous sulfur predictors are developed in past we have developed a sulfur predictor, which estimates feed sulfur concentration before it estimates product sulfur concentration. Our hypothesis is, as product sulfur concentration is affected by feed sulfur concentration, estimating feed sulfur concentration and utilizing that info to predict product sulfur concentration will ensure that our product sulfur estimation is as close as to any given HDS unit (even for online predictions). Therefore, our proposed model will be able to predict product sulfur concentration effectively and continuously. The proposed hybrid model will be validated with lab scale reactor and with industrial data.

A Graphical User Interface (GUI) is design in the later part of this thesis, which will assist implementation of our proposed sulfur predictor on any HDS unit. The sulfur predictor designed in this thesis will be able to overcome all the obstacles mentioned in section 1.3.2 and would be able to predict online product sulfur concentration effectively.

1.5 Thesis Layout

Chapter 1, gives a brief about Hydrodesulfurization and Desulfurization and states the problems encounter when a sulfur predictor is implemented online to estimate product sulfur concentration for HDS unit. It also mentions the objective of this thesis. Chapter 2, states literature review on past developed sulfur predictors for Hydrodesulfurization (HDS) unit. It states about different types of HDS process have been developed to improve the efficiency of desulfurization process. It also covers different types of sulfur predictors developed in past which can capture the trend of hydrodesulfurization process and can predict product sulfur concentration for HDS unit. Chapter 3, states a mathematical model and validates the same for a lab-scale reactor and industrial reactor. The mathematical model validated in this thesis is able to capture the dynamic nature of HDS unit while predicting product sulfur concentration effectively. In this chapter, sensitivity analysis of mathematical model is conducted which justifies the importance of feed sulfur concentration data while predicting product sulfur concentration for HDS unit. In chapter 4, we develop an empirical (data driven) model, which will be part of our proposed hybrid model for online product sulfur concentration prediction. Using Design of experiments technique, we improve the prediction of our data driven model. This new

data driven model is validated and results are stated in this chapter. Chapter 5, states Hybrid model and explains the proposed hybrid model for the HDS unit. This chapter gives details how mathematical model and data driven model developed in earlier chapters are combined together to develop hybrid model, without affecting the speed of simulation. In this chapter, proposed hybrid model validated (with a lab – scale and industrial reactor) and results are stated. In chapter 6, graphical user interfaces (GUI) is design and develop which assists our proposed structure of sulfur model to estimate product sulfur concentration when implemented on an online scenario, for any given HDS unit. In this chapter, GUI is simulated and results are stated using the industrial real time data collected from an HDS unit. Chapter 7 concludes the thesis with conclusion of this thesis.

Chapter 2

2 Literature Survey

2.1 Different types of HDS process

As mentioned in Figure 1.1, there are different ways in which HDS process can be classified. In this section these different types of HDS process are explained in brief as follows

2.1.1 Conventional HDS

Catalytic HDS for crude oil and refinery streams is carried out at elevated temperature. Hydrogen partial pressure converts organosulfur compounds to hydrogen sulfide (H_2S) and hydrocarbons. Detailed description of the HDS process is presented in the literature (Henrik et al., 1996; Chen et al., 2010). The conventional HDS process is usually conducted with $CoMo/Al_2O_3$ and $NiMo/Al_2O_3$ catalysts (Henrik et al., 1996). Factors affecting HDS reactions are desulfurization level, activity and selectivity depends on the properties of the specific catalyst used, the reaction conditions, compositions of various compounds present in the feed stream, and reactor and process design.

2.1.2 HDS by advanced catalyst

Deep desulfurization is achieved by increasing the severity of process conditions. But severe process conditions results not only in deep sulfurization but also undesirable conditions. This could be avoided if catalyst with improved pore activity and selectivity

is used. Advance catalyst like Cobalt – Molybdenum (Co-Mo) and Nickel–Molybdenum (Ni-Mo) with high selectivity for sulfur and nitrogen removal can improve desulfurization process for existing HDS reactors. Also in some cases these catalyst are synthesized which not only have improved pore activity and selectivity but also improve octane number or aromatics content which are essential for high fuel quality. Co-Mo catalyst performs better when operating pressure is low and feed has little amount of impurities. Ni-Mo catalyst performs better for heavier feeds and at higher pressures. The main factors that should be considered in selection of appropriate catalysts for the production of low sulfur fuel are (Vergov et al., 2009; Stratiev et al., 2006).

- The activity of the catalyst in desulfurizing, the least reactive sterically hindered alkyl Dibenzothiophene.
- Feedstock quality, especially the concentration of inhibiting compounds (i.e. hydrogen sulfide and other aromatics) in the feedstock, feed blend components, distillation range, etc.
- Operating variables, especially, the hydrogen partial pressure.

2.1.3 HDS by Advanced Reactor Design

Counter-current operations

Conventionally used hydro treating reactors are fixed – beds with concurrent supply of oil streams and hydrogen, resulting in unfavorable H_2 and H_2S concentration profiles through the reactor. Counter–current operations can provide a more preferable concentration profile. In counter–current reactor operation mode, the oil feed is

introduced into the reactor at the top and hydrogen is introduced at the bottom of the reactor. H_2S is removed from the reactor at the top, avoiding possible recombination of H_2S and olefins at the reactor outlet, this results in better level of desulfurization level as compared to concurrent flow because presence of H_2S inhibits the process. Counter current operation yields lower sulfur, higher conversion and efficient removal of other impurities content in the product even at high velocities when compared to concurrent flow as H_2S formed is removed from the system and have no chance to affect the reaction. (Mederos et al., 2007).

Change of feed quality

The quality of the feedstock plays an important role in hydrodesulfurization (Marafi et al., 2007; Stanislaus et al., 2003). The concentration of nitrogen and presences of other compounds affects the level of desulfurization process. Cracked distillates are more difficult to desulfurize and thus, they should be avoided if existing HDS reactor operates at low pressures. Hydro treating experiments were conducted using feeds under identical conditions in the temperature range 320–380°C. The results showed that the degree of desulfurization of the diesel for the heavy crude was remarkably lower than that of the light crude diesel (Marafi et al., 2007).

Change in feed temperature for the reactor

Reactor temperature is an important variable that has a significant influence on HDS. The start of the run temperature is normally selected based on the required sulfur

reduction which usually depends on the operating conditions. This temperature is gradually increased to maintain the catalyst performance. Studies showed that a 38°C increase in reactor temperature is required to reduce the sulfur content of diesel from 500 to 50 ppm. Further large increases in temperature will be required to reduce the sulfur level to 10 ppm. However, high temperatures increase the rate of catalyst deactivation and reduce the operating range for production of low sulfur concentration fuel (Henrik et al., 1996).

Reduction of liquid hourly space velocity (LHSV) or increase of catalyst volume

Liquid hourly space velocity (LHSV) is the reactant liquid flow rate to the reactor volume at standard temperature. The inverse of LHSV is generally referred as residence time. Decreasing LHSV usually increases the degree of desulfurization and reduces product sulfur concentration due to an increase in the contact time between the liquid feed and the catalyst. Thus by reducing LHSV or increasing catalyst volume the reactor can be operated at lower temperature. Studies have shown that reducing the LHSV by half, results in 20°C decrease in average reactor temperature to achieve the same specific performance when all other operating conditions are unchanged (Knudsen et al., 1999).

All these points are taken into consideration when we design sulfur predictor or HDS unit. Sulfur predictor for HDS reactor in literature can be broadly classified into two categories

- Mathematical Model

- Data Driven Model

2.2 Overview of Mathematical Model developed for HDS system

Knowledge driven parametric models are often referred as “mathematical models”, “mechanistic models” or “first principle model”. Mathematical models are able to capture dynamic behavior of HDS unit while predicting product sulfur concentration (Mederos et al., 2009; Cicili et al., 2009; Deng et al., 2010). For mathematical model, equations are derived based on the following considerations

- Mass and Energy balances.
- System constraints.

For mathematical models, the impact of temperature, feed concentration and effect of any variables on the rate of reaction is reported by kinetic model. The reactor model relates the state variables of the process to other variables. A certain degree of extrapolation beyond the domain of known process operating conditions is one of the greatest strengths of these models (Bhutani et al., 2006). Mathematical models are not as simple to obtain as parametric or data driven models, and owing to the extreme non-linear behavior of HDS it can be difficult. In order to reduce some of complexity, assumptions are generally made, and an easy to understand model are develop which does not appropriately describe the process but sometime still preferred over a very detailed one (Khalfalla et al., 2009; Slimani et al., 2013). The complex mathematical model developed is usually solved by numerical methods.

With increase in demand of low sulfur concentration fuel, studies are focused on HDS unit, where mathematical models were used as HDS equivalent to study effect of various parameters on product sulfur concentration. The parameters studied were, catalyst type: its length and arrangement, (Fujikawa et al., 2009; Oyama et al., 2009), process parameters, feed source and its quality or composition (De la Paz-Zavala et al., 2013; Mochida et al., 1996), inhibitions caused by presence of H_2S (Dong et al., 2014) and reactivity of sulfur (Stanislaus et al., 2003). Mathematical model were used to estimate the kinetic parameters and optimum operating conditions (Jarullah et al., 2011), which would ensure lower sulfur concentration for the fuel.

There are different mathematical models developed in past, some of which are described below:

Avraam et al., 2003, presented a steady–state model for three phase trickle flow fixed bed reactors applied to hydrodesulfurization of light oil feedstocks containing volatile compounds. Mass balances, energy balances and overall two phase flow momentum balance were considered in detail. All physical and chemical properties were estimated as a function of process conditions. The chemical reaction rate was described by Langmuir–Hinshelwood mechanisms where inhibition by H_2S were considered. A collocation method for solving extended system of nonlinear equations was used successfully. The experimental data obtained from a pilot plant unit showed excellent agreement with the mathematical model employed.

Bhaskar et al., 2004, developed a gas–liquid–solid model to simulate the performance of pilot plant and commercial trickle bed reactors applied to the hydro treating of diesel. A three phase heterogeneous model was used and based on two–film theory, where mass transfer phenomena at gas–liquid and liquid–solid interfaces was incorporated. Their model was developed to calculate kinetic parameters HDS reactions using the results obtained in lab scale experiments. A Runge-Kutta fourth order numerical technique was used to integrate the differential equations along the catalyst bed length. The Newton-Raphson method was used to solve the nonlinear equations concerning the surface concentrations of various compounds. The model simulations showed good agreement with the experimental results. The model was also employed to investigate the effect of operating conditions on product quality.

Alvarez et al., 2009 presented a three phase heterogeneous plug flow reactor model for describing the behavior of HDS in a multi–fixed bed reactor system. The main reactions were mass transfer phenomena between gas–liquid and liquid–solid were taken into account. Langmuir–Hinshelwood kinetics were used to model HDS reactions. In order to calculate kinetic parameters, a set of experiments were conducted in a multi reactor pilot plant at several operating conditions. The optimal set of kinetic parameters was calculated using the Levenberg-Marquardt algorithm. The predicted model was found to be satisfactory with experimental results obtained from the lab–scale.

Bahzad et al., 2009, developed a mathematical model to simulate the performance of a pilot plant reactor during hydrodesulfurization reaction. They found some discrepancies between predictions and pilot plant data that may be attributed to the type of catalyst used. In literature, an one-dimensional, plug-flow trickle bed reactor model was developed by Cheng (Cheng et al., 2004) for modelling and simulation of a steady state, adiabatic commercial HDS reactor. The effect of operating conditions on impurities removals were investigated and mentioned in Cheng's report.

Farahani et al., 2011, discussed the process of hydrodesulfurization of heavy gas oil. A non-isothermal heterogeneous reactor model was developed and then simulated with a three stage trickle bed reactor and kinetic models based on the Langmuir-Hingshelwood approach. The results showed good agreement between the simulation results and the experimental data.

Even though these models are able to estimate the product sulfur concentration, mathematical models when implemented on an online system fails to predict product sulfur concentration for HDS unit (Von Stosch et al., 2014). The reason behind this unexpected behavior is that mathematical model needs various parameters, which may or may not be available / recorded for an online system or only partial knowledge about the physical phenomena taking place in the process is available. Mathematical model requires feed sulfur concentration data (which is not recorded for the online system) resulting, in poor prediction of product sulfur concentration. The problem was not only

limited to non-availability of data, mathematical model is calculation intensive. Solving of these numerous correlation and differential equations need time which slows the predictions for product sulfur concentration for an online system. Thus, slow and poor response of mathematical model limits its use as control tool for on-line application in industrial practice (Stanislaus et al., 2003; Bhutani et al., 2006).

2.3 Data Driven Model developed for HDS systems

A data driven model is also known as “regression model”, “empirical model” or “data driven soft sensor”. Data Driven model is a predictive model that describes the relationship between the predicted process variables and the measured variables. Data-driven soft sensors can be applied to the online estimation of product using process measurement data because they have become widely available in many chemical plants (Kadlec et al., 2009). Unlike physical sensors, which directly measure the value of a variable, data driven models measure the process variables. Therefore, data driven model use frequently sampled process variables such as temperature, pressure, flowrate, etc. to measure hard–to–measure variables. There are various modelling techniques to develop data driven models. The most popular modelling techniques used to develop data-driven models are the principal component analysis (PCA) (Jolliffe, 2005), partial least-squares (PLS) (Ahmed et al., 2009; Sharmin et al., 2006), artificial neural networks (Gonzaga et al., 2009; Lou et al., 2012; Mat Noor et al., 2010; Shi et al., 2005; Zhang et al., 2006) and support vector regression (SVR).

Various data driven models were developed in past for Hydrodesulfurization unit. Some of them are stated here;

PLS modelling technique was used to develop data driven model for HDS unit by Qin (Qin et al., 1998).In his work, the recursive PLS algorithm is extended to works block-wise and is thus suitable for adaptive modelling. The algorithm is combined with the two common techniques for adaptive modelling, namely with the moving window and the forgetting factor approaches. The performance of the proposed algorithms is demonstrated by applying it modelling in a refinery process. One of the contributions of this work is the definition of Neural Network Partial Least Squares (NNPLS) algorithm which is a hybrid system combining the PLS algorithm with an MLP. This algorithm makes use of the capabilities of the MLP to map the input variables non-linearly onto the latent variables of the PLS.

Jolliffe 2005, in his work, used PCA modelling technique to develop a data driven model to predict SO_x. A case study applying the proposed method to process industry problems, namely SO_x prediction is discussed. Within the proposed development procedure, firstly missing values are handled using a heuristic approach. This is followed by outlier detection using univariate Hampel identifier and multivariate robust statistics. After the data pre-processing, a PCA data driven model is used to predict SO_x which was within error limits of case study data.

Artificial Neural Network (ANN) modelling technique was used by Berger (Berger et al., 2006) to develop a data driven model for hydrodesulfurization in a mini-pilot plant trickle-bed reactor. Product sulfur concentration was predicted as a function of temperature, pressure, LHSV and staging. The hidden layer contained three neurons. Inputs were normalized in order to give equal importance to each input and to reduce the effect of outliers in the database. The database containing 25 examples which was randomly divided into learn and test sets containing 17 and 8 examples respectively. Almost linear dependence was observed for product sulfur as function of LHSV, however, this behavior does not correspond to experimental data which, indicated that it is necessary to input more data to the model in order to do better learning at low spaces velocities.

Lukec et al., 2008, developed ANN data driven model to determine sulfur content in the hydro-treatment product of distillates. The models were trained using the process and laboratory data from a petroleum refinery. As the models showed to be simple, easy to use, with a good predictability they were used in practice for accurate continuous process monitoring, process fault detection, pointing out a measurement error to the analyzer hardware, estimation of unmeasured states and parameters, process regulation, adaptive control, optimization, and efficient product quality control. This work emphasizes the main advantage of the neural network models that they can estimate the kinetic parameters for different feed conditions.

Behnasr et al., 2015, used support vector regression (SVR) modelling technique to develop a data driven model for sulfur content prediction in HDS process. Due to the huge size of industrial data sets used by soft sensors, training and validation time of SVR model has become a challenging issue. In their study, an accurate and reliable data-driven model was developed by means of a SVR integrated with a data compression technique (VQ/PCA) to predict the sulfur content in an industrial HDS process. The proposed integrated technique incorporated two stages: 1) the data compression stage and 2) the prediction stage. First, the PCA and VQ were applied to reduce the dimensionality of the dataset and then a SVR model was developed. In order to evaluate the performance, a wide range of experimental data according to real condition of the refineries were taken from a HDS setup. Therefore, the results can be generalizable to real processes. The obtained results show that integrated technique (VQ-SVR) was better than (PCA-SVR) in prediction accuracy.

For product sulfur predictions, compared to mathematical model, regression model are easy to implement and faster in prediction. However, these models fail to predict sulfur concentration whenever the feed concentration and operating conditions change beyond ranges of their training data sets. Thus, these data driven sulfur predictors are not able to extrapolate any feed conditions for which they are not trained for. In an ideal scenario, data driven model should be trained and developed using entire historical data of plant which is not feasible. In industries recorded data sets are incomplete, does not cover all operating conditions thus its limits the training and in turn the prediction ability of these

data driven models (Kadlec et al., 2009). Also in some cases some important quality variables are not measured, for example feed sulfur concentration in a HDS process. This further limits the effectiveness of these predictors.

Chapter 3

3 Mechanistic Model

3.1 Literature Review

A Mechanistic model (also known as first principle model (FPM), mathematical model, deterministic model) is used when complete information of the process is accessible and the governing equations of the system are solved using analytical or numerical techniques. Mechanistic model inherits a high degree of knowledge abstraction as it is based on mass balance, energy balance and reaction kinetics (Thompson et al., 1994; Estrada et al., 2006). The advantage of mechanistic models is that these models have a certain degree of extrapolation beyond the domain of the known process operating conditions. Various mechanistic models described in the literature have been used to develop, to capture and to optimize HDS process (Rodriguez et al., 2004).

3.2 Mathematical model

The mathematical model is an offline component of the hybrid model. The mathematical model used for our hybrid model is based on the work of Ancheyta (Ancheyta et al., 2006). This particular mathematical model is selected because Ancheyta and his team developed a mathematical model accurate in predicting product sulfur concentration while capturing the dynamic behavior of the system. Ancheyta had used this mathematical model extensively to study the HDS system and effects of various parameters on product sulfur concentration (Ancheyta et al., 2006; Mederos et al., 2007).

Thus, the mathematical model is able to process disturbance while capturing the dynamic behavior essential for our proposed online hybrid model predictions.

For an HDS reaction, all three phases are present in the reactor. Hence, a hydrodesulfurization (HDS) reaction prefers a trickle bed reactor. A brief description of the chemical reaction, for an HDS unit, is stated in section 1.3. The reactor model considers that HDS reactions take place on the catalyst surface. The concentration profile of reactants and products in a trickle bed reactor model is shown schematically in Figure 3.1.

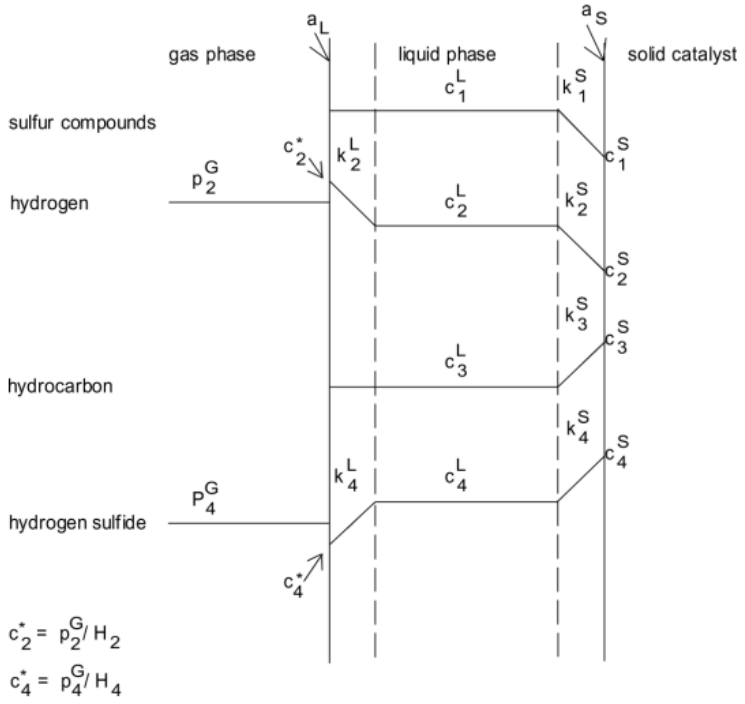

Figure 3.1: Two film model for transport phenomenon of HDS reactor (Shokri et al.,2006)

Figure 3.1 represents two-film transport phenomenon, which is valid for the mass transfer for HDS reaction in the trickle bed reactor. C^i represents the component concentration in each phase. a_l is the liquid-gas mass transfer film cross-sectional area and a_s is the liquid-solid mass transfer film cross-sectional area. k^l, k^s is the mass transfer coefficients for liquid and solid films respectively.

Inside the HDS unit, hydrogen gas (a constituent of the gas–phase) has to be transferred to the liquid phase and then to the catalyst surface in order to react with other reactants. Sulfur has to transfer from the liquid phase to the catalyst surface in order to react with the hydrogen gas. The hydrogen sulfide (H_2S) formed on the solid surface of the catalyst is transferred to the liquid phase and then subsequently to the gas phase. However, there is a probability of H_2S being adsorbed on the catalyst sites inhibiting HDS reactions (Bhaskar et al., 2004).

The assumptions assumed for the HDS unit is as follows

- The reactor is adiabatic
- There are no radial concentration gradients
- The Fick's Law describes the mass transfer.
- The catalyst activity does not change with time.
- Vaporization and condensation do not take place in the reactor.
- There is a negligible pressure drop across the reactor.

- The chemical reaction takes place only at the surface of the catalyst and not in any other phase.

We briefly describe the model for completeness based on these assumptions and chemical reaction knowledge

The mass balance equation (3.1) is the balance equation for the gas phase in the reactor

$$\frac{\epsilon_g}{RT} * \frac{\partial p_{i_g}}{\partial t} = -u_g * \frac{\partial p_{i_g}}{\partial z} * k_{i_l} a_l \left(\frac{p_{i_g}}{H_i} - C_{i_l} \right) \quad \ldots (3.1)$$

where i = H_2 and H_2S

The differential equation of mass balance for the concentrations of hydrogen and hydrogen sulfide in the liquid phase is stated as equation (3.2) by equating the concentrations gradient to the mass transfer of H_2 and H_2S across the gas–liquid and liquid–solid as follows (Bhaskar et al., 2004; Mederos et al., 2007)

$$\epsilon_l * \frac{\partial C_{i_l}}{\partial t} = -u_l * \frac{\partial C_{i_g}}{\partial z} + k_{i_l} a_l \left(\frac{\mathrm{p}_{i_g}}{H_i} - C_{i_l} \right) - k_{i_s} a_s \left(C_{i_l} - C_{i_s} \right) \quad \ldots (3.2)$$

where i = H_2 and H_2S

The mass balance equation for the sulfur present in the liquid phase is stated as equation (3.3) by equating sulfur's liquid–phase concentration gradient to the mass transfer between the liquid–phase and the solid-phase. (Mederos et al., 2007; Alvarez et al., 2009):

$$\epsilon_l * \frac{\partial C_{i_l}}{\partial t} = -u_l * \frac{\partial C_{i_g}}{\partial z} - k_{i_s} a_s \left(C_{i_l} - C_{i_s}\right)$$

…(3.3)

where i = sulfur

The solution of the above equations requires surface concentrations of H_2, H_2S and sulfur. At steady–state, the compounds transported between the liquid phase and the solid phase (on the surface of the catalyst) are consumed or produced through the chemical reaction. By equating the liquid–solid interfacial mass transfer of H_2, H_2S and sulfur components with their reaction rates, we get the following equation (Chowdhury et al., 2002; Bhaskar et al. 2004; Rodriguez et al., 2004; Alvarez et al., 2009):

$$\epsilon_p (1-\epsilon) \frac{\partial C_{i_s}}{\partial t} = k_{i_s} a_s \left(C_{i_l} - C_{i_s}\right) \pm \rho_B \tau \eta_j r_j \left(C_{i_s}, \ldots\ldots, T_s\right)$$

…(3.4)

where i = H_2, sulfur and H_2S

The "-" sign is for the reactants, and the "+" sign is for the products.

We include the energy balance equation as the reactions are exothermic and the system is adiabatic. Energy balance equations for liquid and solid phases are stated as equation (3.5) and equation (3.6) respectively.

For liquid phase

$$\epsilon_l * \frac{\partial T_l}{\partial t} * \rho_l * C_{p_l} = -u_l * \frac{\partial T_l}{\partial z} * \rho_l * C_{p_l} - h_{l_s} a_s (T_l - T_s)$$

…(3.5)

For solid phase

$$(1-\epsilon)*\rho_s*C_{p_s}*\frac{\partial T_s}{\partial t} = h_{l_s}a_s(T_l - T_s) + \sum \rho_B \eta_j r_j\left(C_{i_s}, \dots\dots, T_s\right)\left(-\Delta H_{R_j}\right)$$

…(3.6)

As the heat capacity of the gas is lower to the liquid and solid phases, we do not consider the gas phase energy balance equation (Ancheyta et al., 2006). Developing a kinetic model for HDS reaction is not a simple task because of the complexities of crude oil composition and its analysis. Impurities in crude oil are found in various forms which include but not limited to mercaptan, sulfide, thiophene, dibenzothiophene and their alkyl derivatives shapes. Each form has its own reactivities and a complex reactions ways. For a complex feed, the rate of chemical reaction is usually lumped into a single power law reaction. Langmuir-Hinshelwood models are used for such reactions that include the effect of inhibiting species like H_2S. The hydrodesulfurization reaction is described by the following kinetic equation of the Langmuir–Hinshelwood model that accounts for hydrogen sulfide-inhibiting influence (Bhaskar et al., 2004; Alvarez et al.,2009).

$$r_{HDS} = K_{HDS} * \frac{(C_s)\left(C_{H_2}\right)^{0.45}}{\left(1 + K_{H_2S} * C_{H_2S}\right)^2}$$

…(3.7)

The equilibrium constant of hydrogen sulfide (H_2S) can be described by the Van't Hoff equation (Mederos et al., 2007)

$$K_{H_2S(T)} = 41769.8411\, exp\left(\frac{2761}{RT}\right)$$

…(3.8)

The initial conditions for the above mentioned mathematical model is stated as follows (Mederos et al., 2007)

For t = 0, at z = 0, $(p_{H_2}^g)$= $(p_{H_2}^g)_0$

p_i^g = 0, i= H_2S

c_i^L = $(c_i^L)_0$, i= H_2, S

c_i^L=0, i= H_2S

c_i^S=0, i = H_2, H_2S, S

T=T_0

at z > 0 , p_i^g = 0, i= H_2, H_2S

c_i^L=0, i= H_2, H_2S, S

c_i^S=0, i = H_2, H_2S, S

T=T_0

For the above equations, as mentioned in Nomenclature the symbols $\epsilon, \epsilon_g, \epsilon_l$ and $\epsilon_p$are different notations. The calculations of hold up phase for gas (ϵ_g) and liquid (ϵ_l) is mentioned in study presented by Mederos et al., 2007 and Jarullah et al., 2011.

The oil density as a function of temperature and pressure can be estimated by the Standing–Katz equation, (Macı´as et al., 2004)

$$\rho_L(P,T) = \rho_0 + \Delta\rho_P - \Delta\rho_T$$

$$\Delta \rho_P = [0.167 + (16.181\, x\, 10^{-0.425\, \rho_0})] (\frac{P}{1000}) - 0.01[0.299 + (263\, x\, 10^{-0.0603\, \rho_0})](\frac{P}{1000})^2$$

$$\Delta \rho_T = 0.0133 + 152.4\, (\rho_0 + \Delta \rho_P)^{-2.45}] (T - 520) - [8.1x10^{-6} - 0.0622x10^{-0.764(\rho_0 + \Delta \rho_P)}] (T - 520)^2$$

Henry's coefficients for H_2 and H_2S can be calculated from solubility coefficients (Mederos et al., 2007; Alvarez et al., 2009)

$$H_i = \frac{\upsilon_N}{\lambda_i\, \rho_L}$$

(Korsten et al., 1996) presented the following equations for the solubility of hydrogen and hydrogen sulfide in hydrocarbon mixtures

$$\lambda_{H_2} = -0.559729 - 0.42947x10^{-3}T + 3.07539x10^{-3}(\frac{T}{\rho_{20}}) + 1.94593x10^{-6}T^2 + (\frac{0.835783}{\rho_{20}^2})$$

$$\lambda_{H_2S} = exp\,(3.367 - 0.00847T)$$

The correlations used for estimating the gas–liquid mass transfer coefficients are (Rodriguez et al., 2004; Mederos et al., 2007)

$$\frac{k_{iL} a_L}{D_{iL}} = 7(\frac{G_L}{\mu_L})^{0.4}(\frac{\mu_L}{\rho_L D_{iL}})^{1/2}$$

Glaso's equation, as presented in (Korsten et al., 1996) and (Shokri et al., 2006) is used a generalized mathematical equation for oil viscosity.

$$\mu_L L = 3.141\, x\, 10^{10}(T - 460) - 3.444[log_{10}(API)]^a$$

$$a = 10.313\,[log_{10}(T - 460)] - 36.447$$

The liquid–solid mass transfer coefficients can be calculated from the Van Krevelen Krekels equation as published by various authors (Bhaskar et al., 2004; Mederos et al., 2007),

$$\frac{k_{iS}}{D_{iL} a_S} = 1.8 \left(\frac{G_L}{a_S\, \mu_L}\right)^{1/2} \left(\frac{\mu_L}{\rho_L D_{iL}}\right)^{1/3}$$

In order to determine the liquid – solid and gas – liquid mass transfer coefficients, it is necessary to know the molecular diffusivity of H_2, H_2S and sulfur in liquid. The diffusivity can be calculated by Tyn–calus correlation (Dudukovic et al., 2002, Jarullah et al., 2011)

$$D_{iL} = 8.93\, x\, 10^{-8}(v_L^{0.267} / v_i^{0.433})(T/\mu_L)$$

where i= H_2, H_2S and sulfur in liquid

Other correlations, which are required for simulating this mechanistic model, are stated in below Table 3.1.

Table 3.1: Correlation equations to calculate various parameters for oil and gases (Mederos et al., 2009)

Molar Volume

$$v = 0.285\, v_c^{1.048}$$

$$v_{cm} = 7.5214\, X\, 10^{-3}\, T_{MeABP}^{0.2896}\; d_{15.6}^{-0.7666}$$

Specific Surface Area

$$a_s = (6/d_p)(1 - \epsilon)$$

3.3 Validation and Results

3.3.1 Validation for lab-scale reactor

In order to validate the mechanistic model, we consider a lab –scale reactor and the feed signal mentioned in Table 3.2. We simulate the mechanistic model for the lab –scale reactor with the feed signal as mentioned in Table 3.2 to generate a product sulfur concentration. This product sulfur concentration is referred as simulated product sulfur concentration. Ancehyta uses the data mentioned in Table 3.2 in his work (Ancheyta et al., 2006) to generate the product sulfur concentration for HDS unit (referred as literature product sulfur concentration in this simulation). A comparison plot between two product sulfur concentrations (simulated and literature) is shown in Figure 3.2.

Table 3.2: Operating conditions to simulate a lab – scale reactor

Parameter	Value
Length of Reactor Bed (cm)	31.58
Reactor diameter (cm)	2.54
Temperature (^{0}Celsius)	380
Oil Velocity (cm/sec)	$1.75*10^{-2}$
Gas Velocity (cm/sec)	0.28

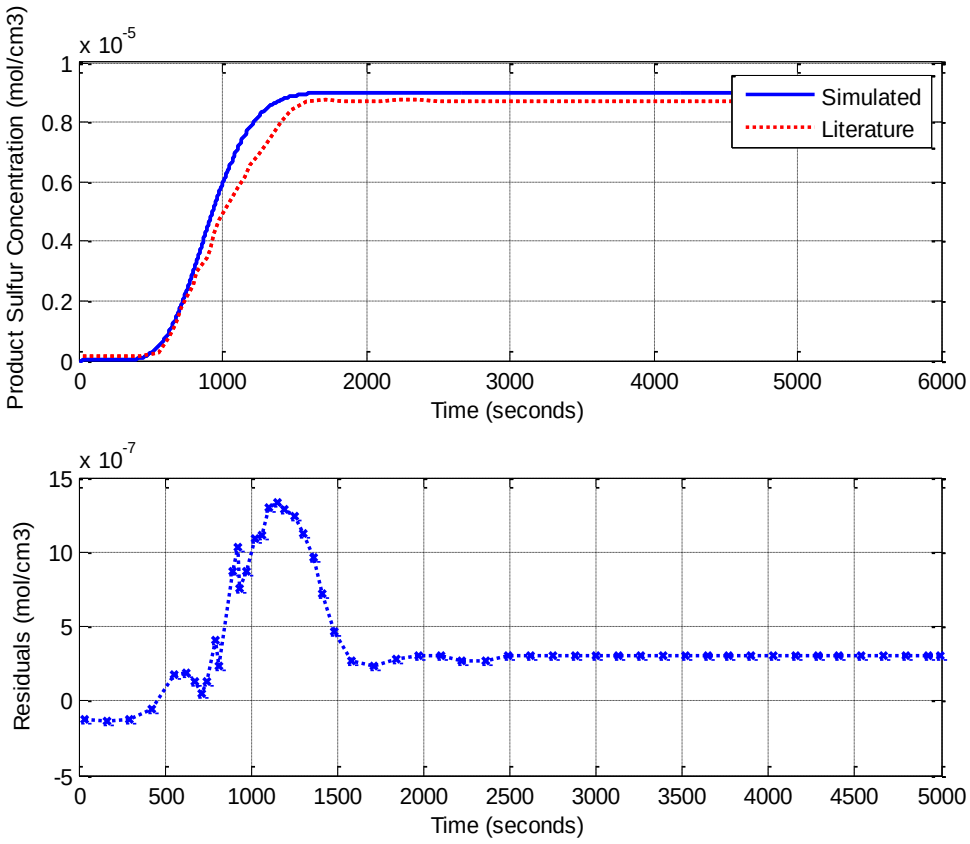

Figure 3.2: Validation of mechanistic model for lab – scale reactor with the data mentioned in (Anchetya et al., 2006)

The residual plot from the above Figure 3.2 indicates that simulation of the mechanistic model is able to predict product sulfur concentration as per the data mentioned in Ancheyta et al.,2006 (mentioned as literature value in the above plot). The comparison plot and the residual plots conclude that results generated by of our mechanistic model can predict product sulfur concentration for a lab – scale reactor with some error when compared to results generated by Ancheyta (Ancheyta et al., 2006).

3.3.2 Validation for Industrial reactor

We consider an industrial reactor, which is stated in Table 3.3 and the operating conditions for our industrial reactor are mentioned in Table 3.3 (Anchetya et al., 2006). After simulating the model using these conditions, we compare our results with the results presented in (Anchetya et al., 2006) paper. The comparison is shown in Figure 3.3.

Table 3.3: Operating conditions to simulate Industrial reactor

Parameter	Value
Length of Reactor Bed (cm)	853.4
Reactor diameter (cm)	304.8
Temperature (^{0}Celsius)	380
Oil Velocity (cm/sec)	0.63
Gas Velocity (cm/sec)	10.27

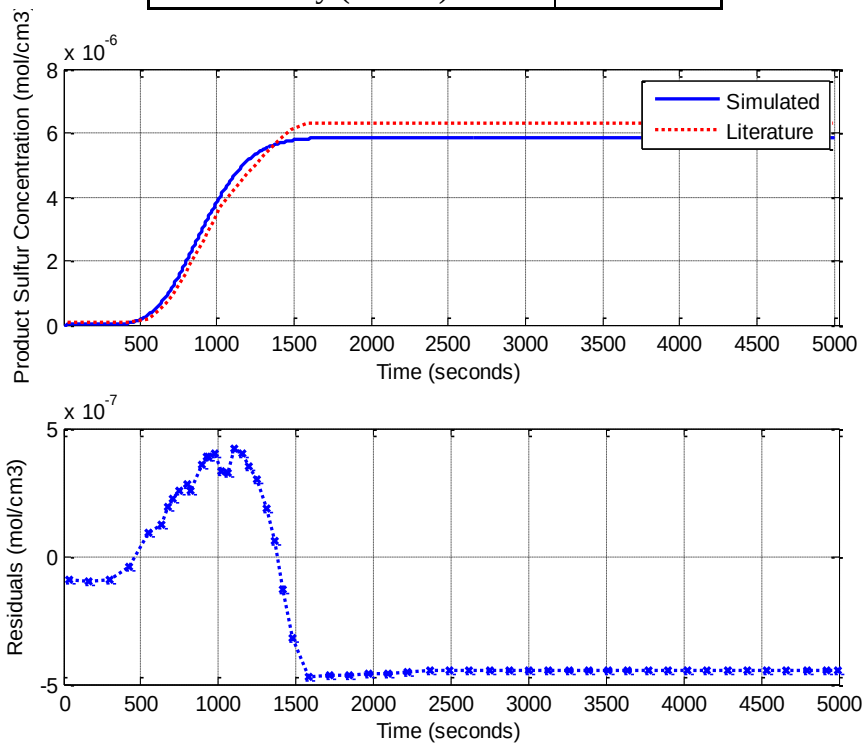

Figure 3.3: Validation of mechanistic model for industrial reactor with the data from (Anchetya et al., 2006)

In the above plot, solid line represents data simulated by the mechanistic model when data from Table 3.3 is simulated. The dotted line represents the data mentioned by (Anchetya et al., 2006) and it is termed as Literature Data in the plot. As we can see from above plot, the mechanistic model can predict the product sulfur concentration satisfactorily as compare to data presented by (Ancheyta et al., 2006) referred as literature data in the plot. The plot indicates that our validation generated results close to the results mentioned by Ancheyta in his work (Ancheyta et al., 2006)

In above Figure 3.2 and Figure 3.3, there is a very small error in our simulated results. This is because our mechanistic model will be a part of the hybrid model for online product sulfur concentration, thus we have reduced the complexity of the mechanistic model. As the error between our simulated results and results stated by (Ancheyta et al., 2006) is minimal, we can conclude that our mechanistic model is able to predict product sulfur concentration for an HDS unit within acceptable error limits. As there are three phases present in the reactor, the feed flowrate and catalyst dimensions are mentioned in Table 3.1 and 3.2 respectively. This is to ensure that the residence time for the reactor is same. If this does not happened to be true for any industrial reactor, the mathematical model will calculate the necessary parameters as per the given feed flowrates and its effect will be taken into consideration while estimating feed sulfur concentration, which will in turn affect the product sulfur concentration prediction.

3.4 Sensitivity Analysis

In an online scenario, the feed concentration tends to fluctuate as effluents from various units are passed through HDS unit. Therefore, in this section, our aim is to study how sensitive our mechanistic model is when subjected to fluctuation in feed parameters. The changes are mentioned in Table 3.4. For our sensitivity analysis, we change one variable at a time (by giving a step change) and study how it affects the product sulfur concentration. The results are simulated and sensitivity analysis plot is shown in Figure 3.4

Table 3.4: Sensitivity Analysis of our Mechanistic Model

Sr. No	Feed Parameters	Nominal Operating Range	Average Amplitude of Step Change	Average Response Time (minutes)	Average Change in Product Sulfur (ppm)
1	Velocity of oil (cm/sec)	$1.729 \times 10^{-2} \pm 25\%$	0.4375×10^{-2}	20-22	1.5
2	Temperature (Kelvin)	600 ± 50	40	20-25	1.9
3	Feed Sulfur Concentration (wt%)	$5\% \pm 2\%$	2	20-25	5

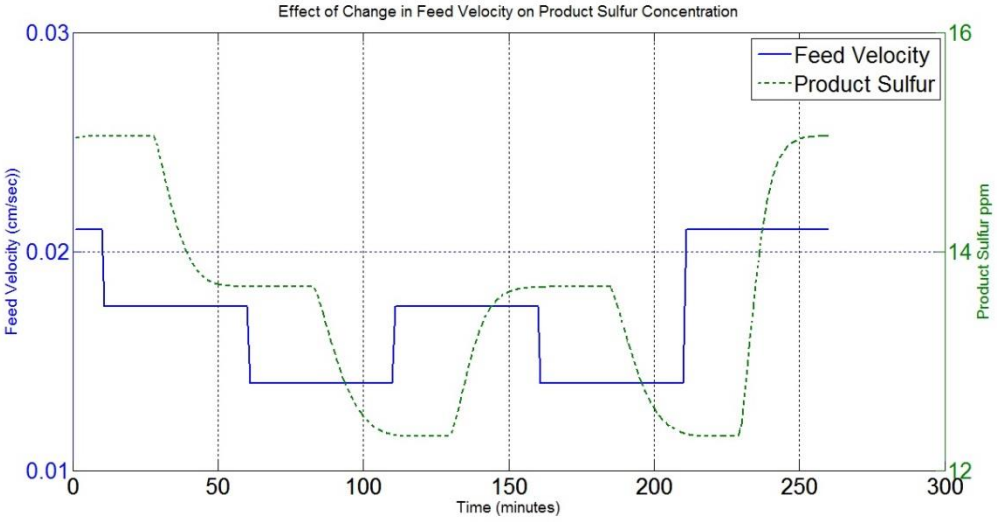

Figure (3.4a): Effect of Feed Oil velocity on Product Sulfur concentration

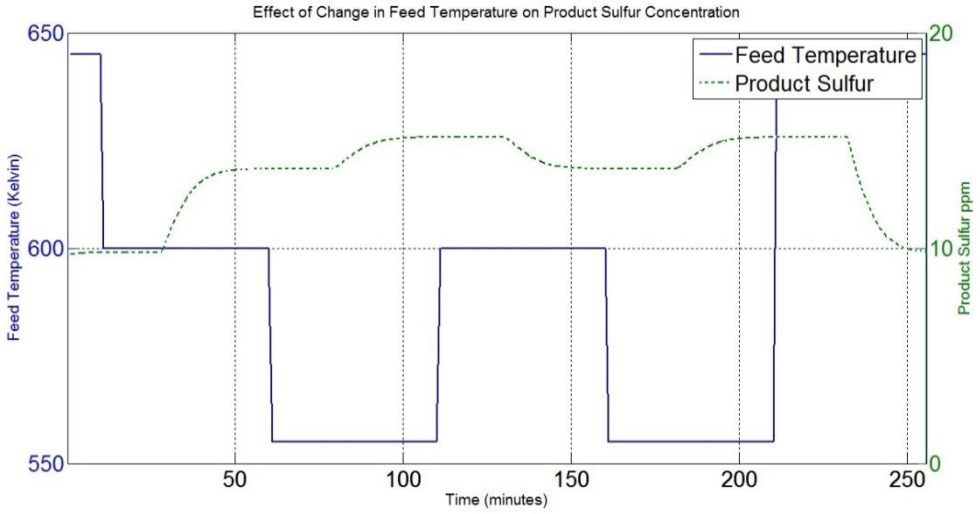

Figure (3.4b): Effect of Feed Temperature on Product Sulfur Concentration

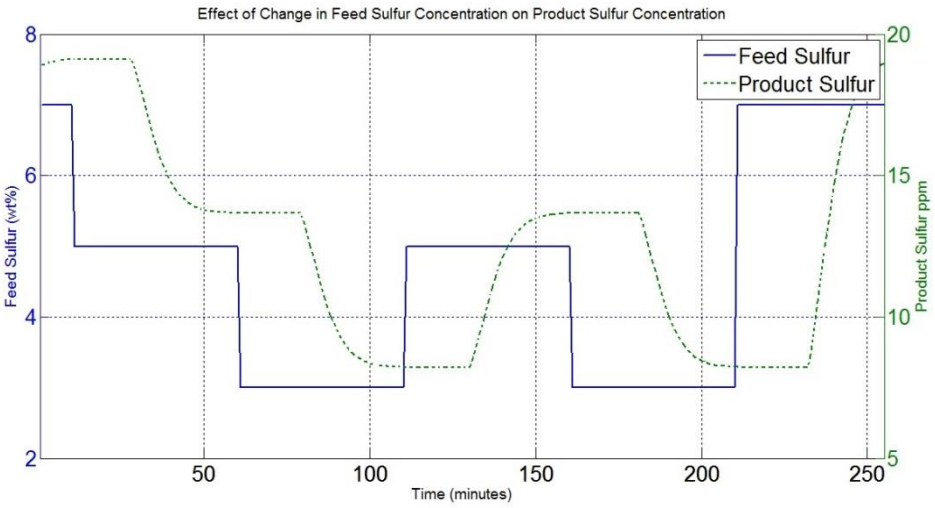

Figure (3.4c): Effect of Feed Sulfur Concentration on Product Sulfur Concentration

Figure 3.4: Sensitivity Analysis of Mechanistic Model

Figure 3.4 shows the sensitivity analysis results for the mechanistic model. In the above plots, the solid lines represent the feed oil velocity (for Figure (3.4a)), feed temperature (Figure (3.4b)) and feed sulfur concentration (Figure (3.4c)), while the dotted line represents the product sulfur concentration for HDS unit (in all the three plots). Figure (3.4a) shows the effect of the change in feed oil velocity on product sulfur concentration. We can conclude from the plot that for our mechanistic model, product sulfur concentration is directly proportional to feed oil velocity. From the Figure (3.4b) concludes that product sulfur concentration is inversely proportional to feed temperature. This is because an increase in reaction temperature can substantially enhance the rate of catalytic reaction and favors the sulfur removal. Figure (3.4c) represents the effect of feed sulfur concentration on product sulfur concentration. It is evident from the plot, product sulfur concentration is directly proportional to feed sulfur concentration. This concludes

that any changes made in feed sulfur concentration will affect the product sulfur concentration.

Thus in presence of fluctuating feed concentration where feed sulfur concentration is not recorded predicting product sulfur concentration will not be accurate. This is one of the reasons, why mechanistic model tends to fail to predict product sulfur concentration when implemented online to predict product sulfur concentration. Also, mechanistic models require time to simulate which is not ideal for online prediction, hence data-driven model are often used for online prediction for HDS unit.

3.5 Conclusions

In this chapter, we have stated a mechanistic model based on the work of Ancheyta, in order to study dynamic behavior for HDS unit. The model stated here comprise mass balance, energy balance and chemical kinetics of HDS process. The mechanistic model is validated for a lab – scale reactor and industrial reactor and found to be predicting the product sulfur concentration as close as to the actual HDS unit. The stated model here is able to capture not only the steady state but also the dynamic state for HDS unit. The sensitivity analysis concluded that mechanistic model is sensitive to feed oil velocity, feed temperature and feed sulfur concentration.

Chapter 4

4 Empirical Model

4.1 Literature Review

Empirical model are easier to develop and implement than mathematical models. These models describe the relationship between the predicted variables and the measured variables (Kadlec et al., 2009). There are different modelling techniques to develop a data driven model or an empirical model for the HDS unit, each with their own advantages and disadvantages. An overview of the various soft sensor modeling techniques had been mentioned in literature (Gonzalez et al., 1999; Sliskovic et al., 2011).

The most popular soft sensors modeling techniques include Principal Component Analysis (PCA) (Jolliffe 2005), Partial Least Square (PLS) (Wold et al., 2001), Artificial Neural Network (ANN) (Principle et al., 2000) and Support Vector Machine (SVM) (Yin et al., 2014). PCA/PLS-based modeling techniques are mostly linear in nature. ANN is a well-established and powerful algorithm for nonlinear modeling, however, it has some limitations. One of them is that ANN-based models are prone to get stuck in local minima during the training and thus may achieve a suboptimal performance on the test data (Xiong et al., 2014). SVM proposed by Vladimir Vapnik (Vapnik et al., 2000) has benefits over other methods, which includes, not being stuck in local minima, less dependence on the amount of training data samples and better generalization ability (Feng et al., 2003; Yan et al., 2004). In literature PLS modeling technique and in recent

time SVR modeling technique is most favored modeling technique used to develop data driven models for HDS unit (Sliskovic et al., 2011; Wold et al., 2001; Yin et al., 2014). A brief about these data driven modelling techniques are stated below.

4.2 Partial Least Square (PLS)

Partial Least Square algorithm, instead of focusing on the covering of the input space variance, pays attention to the covariance matrix that brings together the input and the output data space. The algorithm decomposes the input and output space simultaneously while keeping the orthogonality constraint. In this way, it is assured that the model focuses on the relation between the input and output variables. The general model of PLS is

$$X = TP^T + E$$

$$Y = UQ^T + F$$

where X is n x m matrix of predictors, Y is n x p matrix of response, T and U are n x l matrices, which are projection of X and Y respectively, P and Q are loading matrices of m x l and p x l respectively and E and F are the error terms assumed to be independent and distributed and random normal variables (Wold et al., 2001). It is a modelling technique only for linear relations.

As PLS is a very popular data drive modelling technique in chemical engineering, there are several publications dealing with the application aspects of PLS (Kourti et al., 2002). The original PLS algorithm suffers from similar problems as its PCA counterpart. It is

modelling technique suitable only to linear relations between the data. Therefore, there are some advanced versions of the PLS proposed, which are suited to nonlinear systems. Qin (Qin et al., 1998) proposed an adaptive version of the PLS called recursive PLS (RPLS). Another adaptive version of the PLS based on the moving window technique is the exponentially weighted PLS (EWPLS) (Dayal et al., 1997). There are other variation to PLS modelling technique which includes Neural Network PLS (NNPLS) (Qin et al., 1998), multi – way PLS (Dayal et al., 1997) and exponentially weighted PLS (EWPLS) (Bro et al., 1996)

4.3 Support Vector Regression (SVR)

SVR gained attention within computational learning community because of its strong theoretical foundation in statistical learning theory. The derivation and theoretical justification were stated by Vapnik (Vapnik et al., 2000) while application aspects of SVR briefly discussed by other authors (Rodriguez et al., 2004; Feng et al., 2003). SVR utilizes a nonlinear relation of the form y = f(x) between the vectors of observation x and the desired y from a given set of training samples. A number of cost functions such as the Laplacian, Huber's, Gaussian, and ε-insensitive can be used for the SVR formulation. Among these, the robust ε-insensitive loss function (L_ϵ) is more common (Xiong et al., 2014).

$$L_\epsilon(f(x) - y) = \begin{cases} |f(x) - y| - \epsilon \; for \; |f(x) - y| \geq \epsilon \\ 0 \qquad\qquad\qquad\qquad otherwise \end{cases}$$

where ε is a precision parameter representing the radius of the tube located around the regression function, f(x) (Figure 4.1). The region within this tube is called ε-

insensitive zone. The algorithm tries to minimize the error outside this zone while keeping the complexity of the model to minimal. c and ε are user-defined parameters in the empirical analysis where c > 0. Penalty parameter determining the trade-off between generalization ability and accuracy in the training data, while the parameter ε defines the degree of tolerance to errors. The optimization problem can be reformulated as:

$$\text{minimize} \quad \frac{1}{2}\|w\|^2 + c\sum_{i=1}^{n}(\xi_i^+ + \xi_i^-)$$

This is subject to the conditions

$$y = \begin{cases} y_i - (\langle w, x_i\rangle + b) \le \epsilon + \xi_i \\ (\langle w, x_i\rangle + b) - y_i \le \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases}$$

The positive slack variables ξ and ξ^*represent the distance from actual values to the corresponding boundary values of the ε-tube.

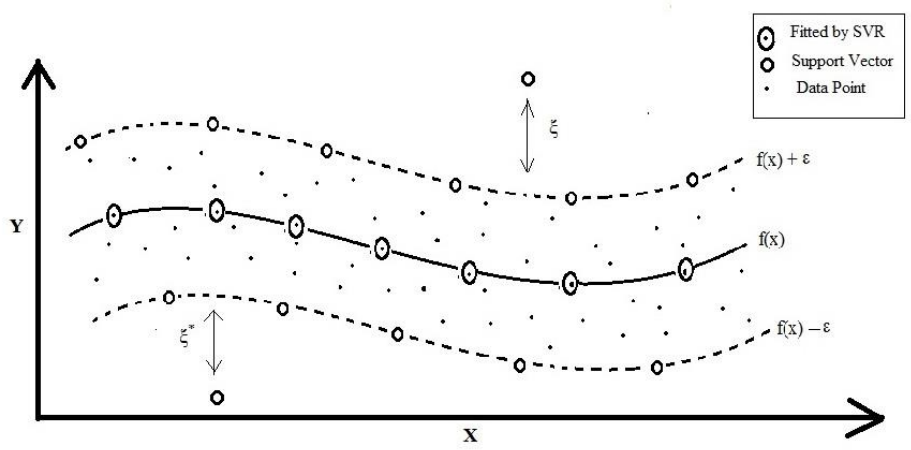

Figure 4.1: A schematic diagram of SVR using ϵ sensitive loss function (Shokri et al., 2006)

While grounded in the theory, SVRs have been demonstrated to work very well for a wide spectrum of applications so it is not surprising that they have also been successfully applied as Soft Sensors. In recent time SVR, regression is often used to develop a data driven model as these models gives a better prediction when compared to other models. (Shokri et al., 2006)

In order to decide which of these data driven modelling technique (PLS or SVR) will be better suited for proposed hybrid model, we compare these models. The comparison is stated in the Figure 4.4.

4.4 CCD Design - Design of experiment technique

SVR, despite being an effective regression model, suffers from the same limitations of other data- driven models (Montgomery et al., 2010). Unless the training data set covers the range well, the model fails to predict in that region. In order to train this model effectively, we have designed a training data set using CCD - Design of Experiments technique to train the SVR model.

4.4.1 Central Composite Design (CCD) theory

The Design of Experiments (DoE) states studying one variable at a time (OVAT) is not an ideal way to study a system (Laird et al., 2009). DoE suggests an alternative approach of full/fractional design or response surface methodology (RSM) (Box et al., 1951).

These alternative approaches are a dynamic, continuous learning process when compared to OVAT. One of DoE modeling technique is Central Composite Design (CCD). As, CCD suggest an additional data points / levels within the given data set, this generates a better model which increase accuracy in predictions (Desai et al., 2006).

A central composite design is a 2k full factorial design to which central point and star points are added (Lye et al., 2013; Montgomery 2010). The star points are the sample points in which all parameters but one are set at the mean level 'm'. The value of remaining parameters is given in terms of distance from the central point. The distance between the central point and each full factorial point can be normalized to 1. There are few types of CCD designs, which depends on the distance between the points and are shown in Figure 4.2

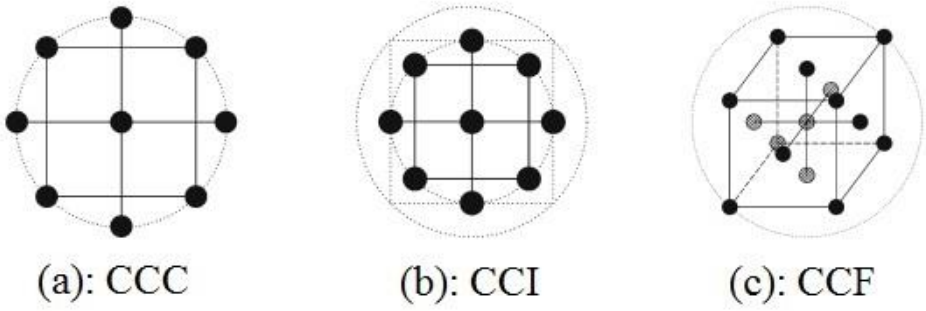

Figure 4.2: Different types of CCD design

Based on the Figure 4.2 various types of CCD design model is explain (Lye et al., 2013):

CCC: Central Composite Circumscribed (CCC) designs are the original form of the central composite design. The star points are at $\sqrt{\alpha}$ = 1, from the center based on the properties desired for the design and the number of factors in the design. These designs have circular, spherical, or hyper-spherical symmetry and require 5 levels for each factor. Augmenting an existing factorial or resolution 5 fractional factorial design with star points can produce this design.

CCI: Central Composite Inscribed (CCI) design, for those situations in which the limits specified for factor settings are truly limits. The CCI design uses the factor settings as the star points and creates a factorial or fractional factorial design within those limits. In other words, a CCI design is a scaled down CCC design with each factor level of the CCC design divided by $\sqrt{\alpha}$ to generate the CCI design). This design also requires 5 levels of each factor.

CCF: Central Composite Face – Centered (CCF) design, in this design the star points are at the center of each face of the factorial space, so $\alpha = \pm 1$. This variety requires 3 levels of each factor. Augmenting an existing factorial or resolution 5 design with appropriate star points can also produce this design

In our case, as we are handling the historical data, the limits specified for factor ranges are true limits, therefore we use CCI version of CCD design. The experimental design structure for CCD model is stated in Table 4.1.

Table 4.1: Localized CCD Design Model for HDS unit.

	Independent Variable		
No of runs	Feed Oil Velocity	Feed Temperature	Feed Sulfur Concentration
1	-1	-1	-1
2	1	-1	-1
3	-1	1	-1
4	1	1	-1
5	-1	-1	1
6	1	-1	1
7	-1	1	1
8	1	1	1
9	-α	0	0
10	+α	0	0
11	0	-α	0
12	0	+α	0
13	0	0	-α
14	0	0	+α
15	0	0	0
16	0	0	0

Using the above scheme (Table 4.1) suggested by the CCD design first SVR model was trained with the available historical data.

In an ideal scenario, to develop a SVR model plant historical data covering the entire region of operation is required, which is rarely available in an industrial scenario. Hence, when only available historical data set was used to develop a model for the entire range of operation, we call this global model, the prediction performance of the global model was not satisfactory for regions where data were scarce. In order to overcome this, for regions where prediction were poor we developed several local CCD models for those

regions. Figure 4.3 explains, the methodology, which is implemented to train and develop the surrogate SVR model. . The local CCD models were used for interpolating the data points. These interpolated points from the local CCD models were included into the available historical data set. Finally, the global SVR model (i.e. surrogate SVR model) is trained and developed based on the updated data set comprised of historical data and interpolated points. This surrogate SVR model has better information about the system and therefore can predict better than previously developed SVR model. Thus, inclusion of additional data points suggested by localized CCD model into available historical data improves the quality of historical data, resulting a better predicting surrogate SVR model.

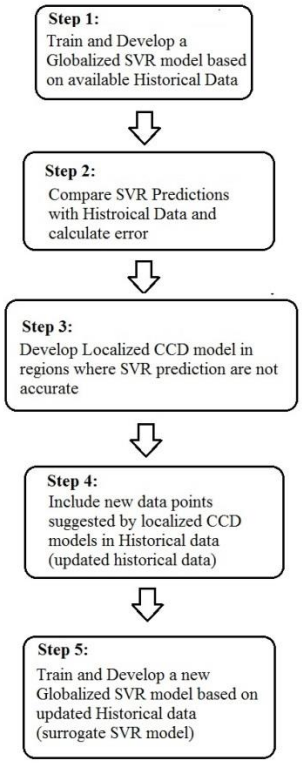

Figure 4.3: Flowchart representing methodology for developing surrogate SVR model

4.5 Validation and Results

4.5.1 Comparison between PLS and SVR empirical modeling technique

We consider a lab-scale reactor as mentioned in Table 3.2. For the lab –scale simulated reactor we simulate our mathematical model with the Feed ranges mentioned in Table 4.2 to predict product sulfur concentration. In analogy to an industrial scenario, we consider, feed velocity of oil, temperature and product sulfur concentration as our historical data.

Table 4.2: Feed Signal to Generate Historical Data for Lab –scale reactor.

Input Variable	Range
Velocity of Oil (cm/sec)	$1.725 \times 10^{-2} \pm 25\%$
Temperature (Kelvin)	600 ± 20
Input Sulfur (wt%)	$5\% \pm 2\%$

Using the historical data, we train and develop PLS and SVR models, a comparison plot between the PLS and SVR model is shown in Figure 4.4

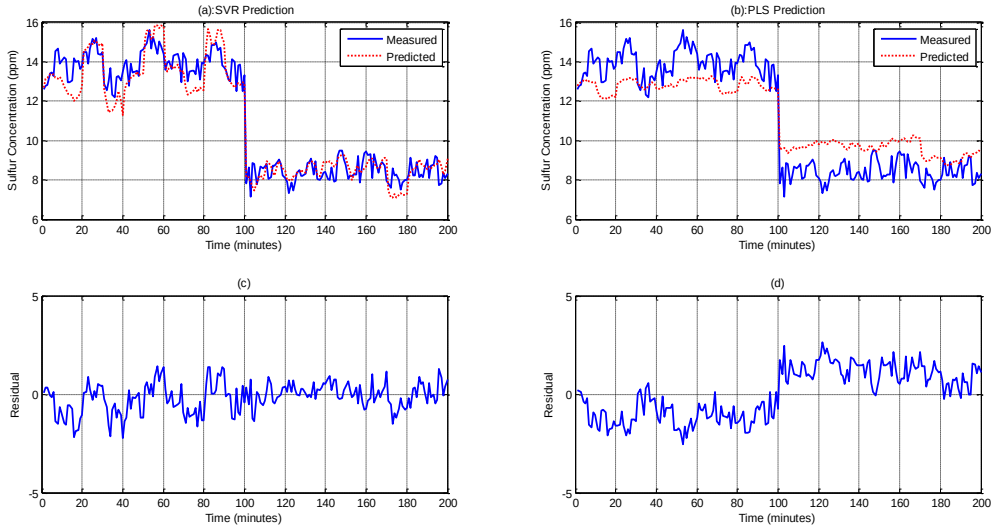

Figure 4.4: Comparison between PLS model and SVR model prediction for HDS unit.

In Figure 4.4, plot (a) represent SVR predictions while plot (b) shows PLS predictions, plot (c) and plot (d) are the residuals plots for SVR and PLS respectively. From the above plot, we can conclude that PLS model can predict product sulfur concentration; however, SVR predictions are better. The same is can also be concluded by residuals plot (subplot (c) and (d)), as magnitudes of SVR residuals are around '2 ppm' and remain to be same after 100^{th} point when product sulfur concentration changes. However for the PLS model predictions, the fluctuations in the residuals increase with the change in product sulfur. Thus, we select we SVR algorithm as our data driven modeling technique.

4.5.2 Validation of Methodology

4.5.2.1 Lab-scale reactor

We improve our SVR data-driven model prediction using the methodology shown in Figure 4.2. The localized CCD models are developed as per the localized CCD design Table 3.4. The independent variables and their coded and actual value ranges for one of the localized model is stated in the following Table 4.3.

Table 4.3: Independent variables and their coded and actual values of a localized CCD model

Independent Variable	Levels ranges				
	-1	- α	0	+ α	+1
Feed Oil Velocity	1.32	1.456	1.785	2.072	2.25
Feed Temperature	580	586	598	609	616
Feed Sulfur Concentration	3.5	3.73	4	4.26	4.5

Localized CCD model is developed using a Design of Experiments (DoE) software for the given dataset. One of the localized equations generated for CCD model is stated below as the following equation (4.1)

$$
\begin{aligned}
&Product\ \ Sulfur\ Concentration \\
&= -2.19098 - 0.11046\ (Feed\ Oil\ Velocity) + 7.691 \\
&* 10^{-3}(Feed\ Tempertaure) + 2.8057\ (Feed\ Sulfur\ Concentration) \\
&+ 1.86 * 10^{-4}(Feed\ Oil\ Velocity) * (Feed\ Temperature) + 1.057 \\
&* 10^{-3}(Feed\ Oil\ Velocity) * (Feed\ Sulfur\ Concentration) - 9.4 \\
&* 10^{-5}(Feed\ Temperature) * (Feed\ \ Sufur\ Concentration) - 6.7 \\
&* 10^{-6}\ (Feed\ Temperature)^{\,2}
\end{aligned}
\qquad \ldots (4.1)
$$

Thus by adding additional points suggested by CCD model into historical data , we develop another SVR model (surrogate SVR model).

For the lab –scale simulated reactor (mentioned in Table 3.2) we simulate our mathematical model with the Feed ranges mentioned in Table 4.2 to predict product sulfur concentration. In analogy to an industrial scenario, we consider, feed velocity of oil, temperature and product sulfur concentration as our historical data.

Using 70% of the historical data we trained SVR model and the remaining data were used to validate the SVR model. The comparison between SVR model predictions with measured sulfur concentration, from historical data is shown in Figure 4.5. In next step, localized CCD models were developed for the entire historical data and the new data points suggested were included in historical data. Thus, we have more data points in our historical data to train our surrogate SVR model. Figure 4.6 shows a comparison of surrogate SVR model (i.e. SVR model developed by incorporating additional CCD suggested data points along with historical data) predictions, with measured product sulfur concentration from historical data.

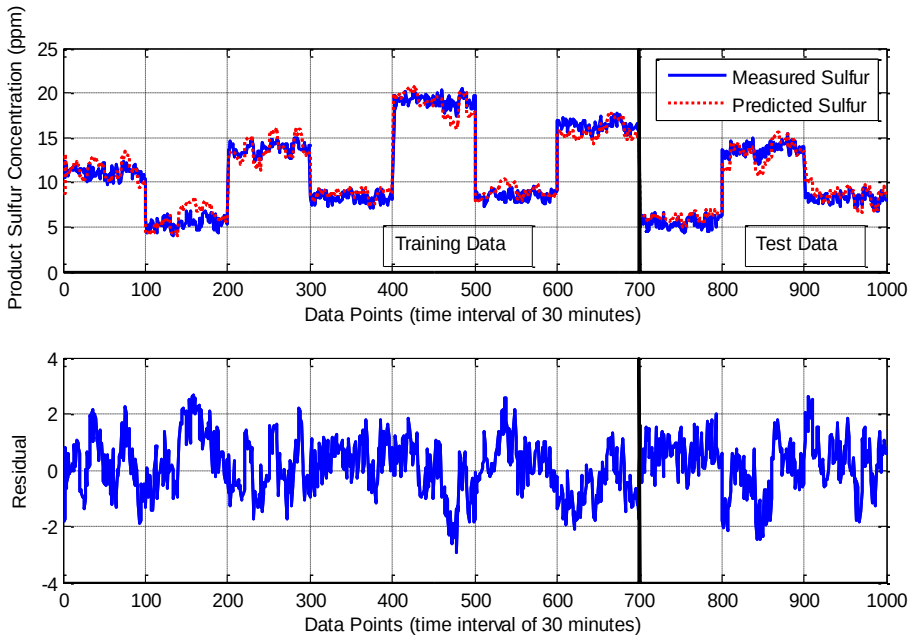

Figure 4.5: Comparison between the SVR prediction and Measured Sulfur concentration.

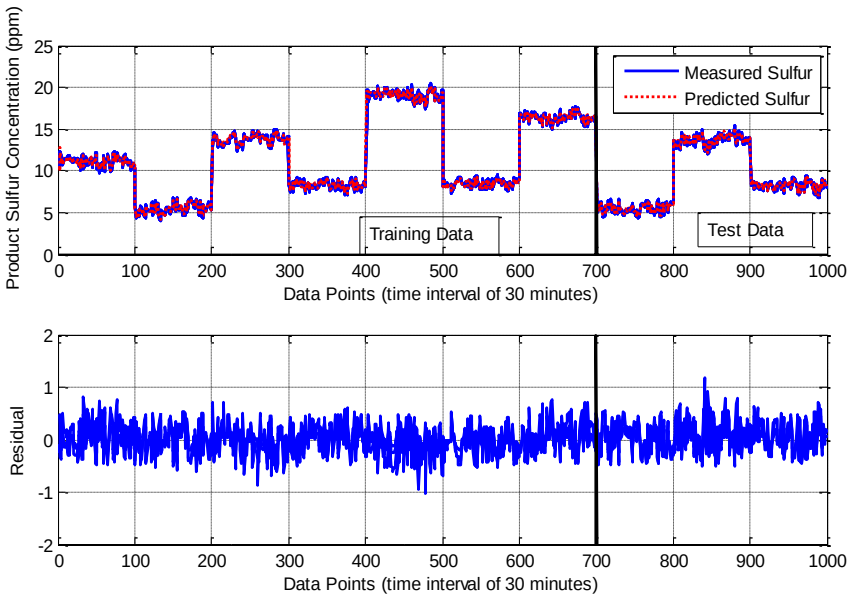

Figure 4.6: Comparison between the surrogate SVR prediction and Measured Sulfur concentration.

Figure 4.5 shows that the SVR model developed on the basis of 'historical data only' is capable to predict the historical product sulfur concentration with an average error of 15% which was consistent with training data set. Figure 4.6 compares the prediction of surrogate SVR model with 'historical data and imputed data from the localized CCD model'. The average error for these predictions, both training and test data set is 10%, which is better than the SVR prediction in Figure 4.5. Thus, as compared to the SVR model developed, surrogate SVR model has more data points to train. By including these imputed data points for training overall prediction capability of the data based model can be improved.

4.5.2.2 For industrial plant data

The specifications and operating ranges limits of the industrial reactor are mentioned in Table 4.4 and 4.5 respectively.

Table 4.4 : Specification of Industrial Reactor.

Feed	
Flowrate	Data stated in MBPD
Temperature	Data stated in ^{0}Fahrenit
LHSV	2 hour^{-1}
Hydrogen to oil ratio	2000 std cubic feet per barrel
Catalyst length	5 m
Diameter	2 m
Number of catalyst beds (in series arrangement)	3
Bulk density	816.3kg/m^{3}
Composition	
Carbon	80 wt%
Hydrogen	15 wt%
Sulfur	5 wt%

Table 4.5: Input parameters and their ranges

Input Parameter	Range
Velocity of Oil (MMBPD)	12.5 – 52.2
Temperature (Kelvin)	520 – 630

It is evident from given industrial data that feed sulfur concentration is not available. Thus, our historical data consists of velocity of oil, temperature and analyzer sulfur concentration for the industrial reactor. The analyzer is located at the tankage, the measured sulfur will be available with a time delay.

We first improve our SVR data-driven model prediction using the methodology mentioned in Figure 4.2. The localized CCD models are developed as per the localized CCD design Table 4.1. The independent variables and their coded and actual value ranges for one of the localized model is stated in the following Table 4.6.

Table 4.6 : Independent variables and their coded and actual values of one localized model

Independent Variable	Levels ranges				
	-1	- α	0	+ α	+1
Feed Oil Velocity	24.64	28.809	35.45	42.67	46.26
Feed Temperature	576	593	622	654	668

We followed the methodology mentioned in Figure 4.2 to develop, train and validate SVR model and the surrogate SVR model. The comparison of SVR model and surrogate SVR model predicted product sulfur concentration with measured sulfur concentration is shown in below figures. Figures 4.7 and 4.8 are SVR model and surrogate SVR model predictions for industrial data respectively. Figure 4.9 compares the mean squared error for SVR and surrogate SVR model for training and validating (test) data sets.

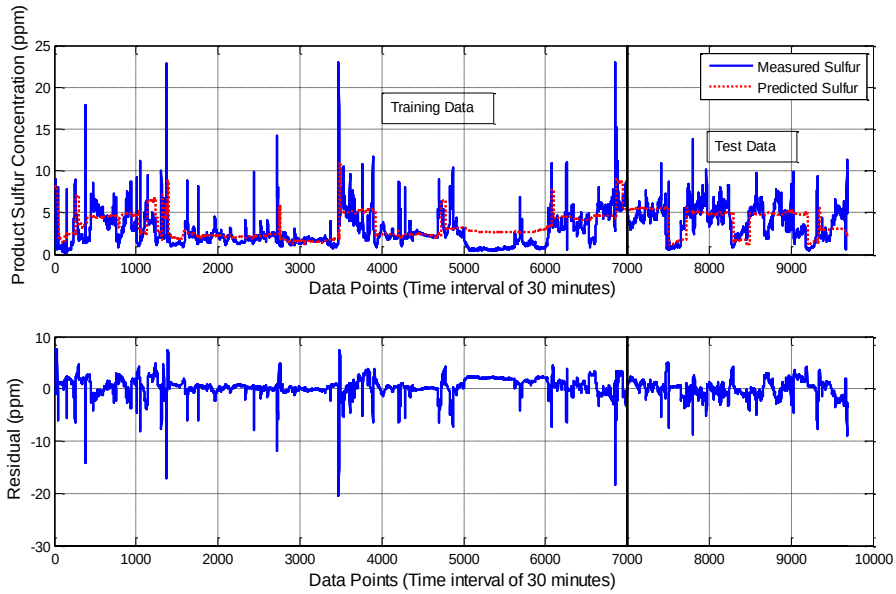

Figure 4.7: Comparison between the SVR prediction and Measured Sulfur concentration.

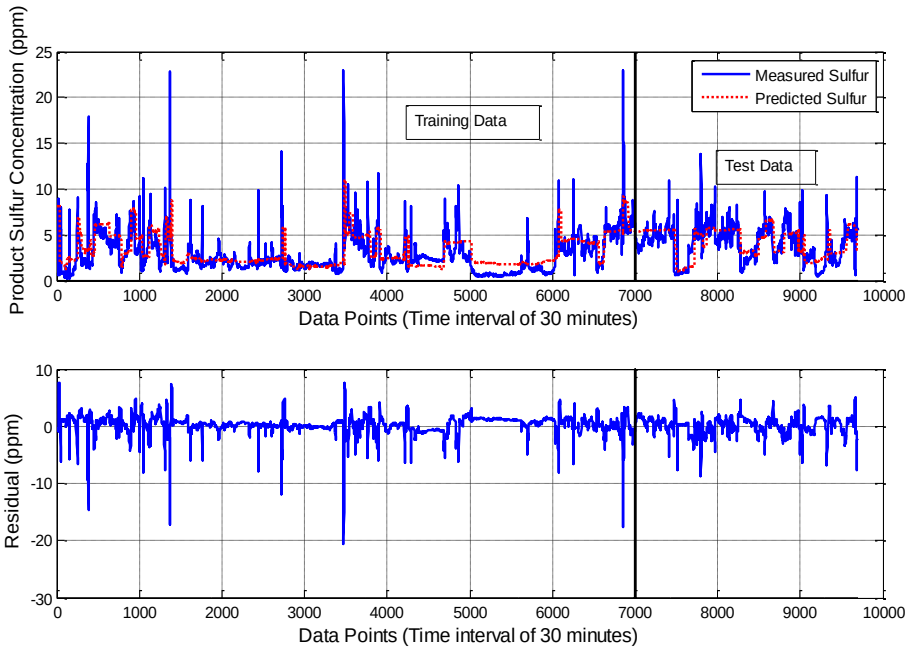

Figure 4.8: Comparison between the surrogate SVR prediction and Measured Sulfur concentration.

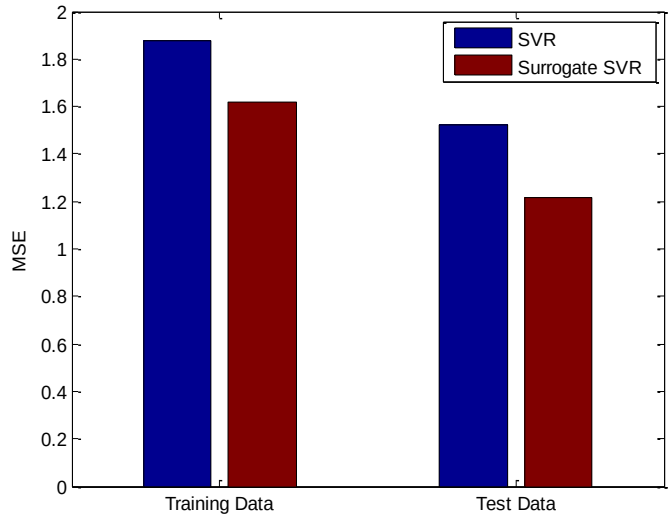

Figure 4.9: Comparison for MSE's for SVR and surrogate SVR model for training and test data.

Comparing Figure 4.7 with Figure 4.8 we conclude that surrogate SVR model predicts with higher accuracy for the industrial data set. In Figure 4.8, the sudden spikes (one or two points) in the plot were considered as disturbances or faulty data recordings, which was part of the measured data. Thus while predicting the product sulfur concentration, our predictor is not able to predict those points resulting an error of range 10 to 25ppm. Figure 4.9 also confirms that the mean squared error for predictions of surrogate SVR model is lower than SVR model, which again confirms that surrogate SVR model predicts better of the two models. Thus, we will be using surrogate SVR data-driven model as our data-driven model in proposed hybrid model.

4.6 Conclusions

In this chapter, we had developed an empirical model / data driven model to be a part of our proposed online Hybrid model for product sulfur concentration of HDS unit. In literature various data driven model are used in past to develop HDS model (Sliskovic et al., 2011). Therefore we compare PLS and SVR model and selected SVR model as its better suitable to capture nonlinear models. In order to improve the predictions of SVR data driven model, we develop a methodology. The methodology is explained in detail in this chapter. The methodology use Design of Experiments – CCD technique. This methodology is validated for both lab-scale reactor and industrial reactor. Thus, the new methodology improves the data driven model predictions, which helps us to develop a better predicting SVR model for proposed hybrid model.

Chapter 5

5 Hybrid model

5.1 Literature Review

Various sulfur predictors (models) are generated and developed in past which are able to predict product sulfur concentration for hydrodesulfurization (HDS) unit (Von Stosch et al., 2014). (S. Porto et al., 2011) in his study of various sulfur predicting model stated that these proposed sulfur models can be classified into three board categories. (i) mathematical or parametric (when process behavior can be represented mathematically with algebraic and/or differential equations), (ii) regression or non-parametric models (when no physical insight is available) and (iii) semi-parametric or hybrid model (when some physical insight is available). The same is shown in Figure 5.1.

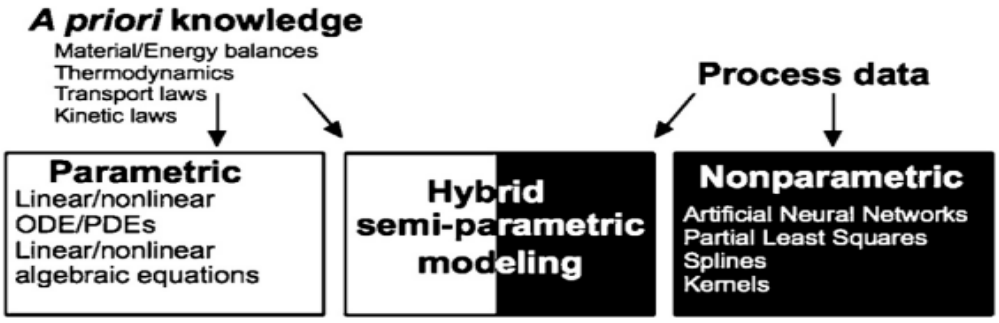

Figure 5.1: Mathematical, Regression and Hybrid modeling based on their types of knowledge.(S.Porto et al., 2011)

When mathematical and data driven model are combined together these models complement each other and can predict better (S. Porto et al., 2011). (Cicili et al., 2009), developed a hybrid model where PLS data driven model was used to estimate the missing data, while mathematical model was used to estimate the output sulfur concentration. The hybrid model was able to predict better and was able to extrapolate and interpolate more accurately than data driven model. Unlike mathematical model, hybrid models do not require a huge amount of initial data, these models are not calculation oriented nor depend solely on the historical data for any predictions (Porto et al., 2011; Bhutani et al., 2006; Gernaey et al., 2010). Hybrid models are also used for HDS modelling which included monitoring and optimization of process conditions. Since, hybrid models perform better than the mathematical and data driven models, these models are more favored as the HDS equivalent models in recent times (Kadlec et al., 2009).

Various hybrid models have been developed in literature of offline implementation for HDS unit. (Behnasr et al., 2015) developed a hybrid model where the data is filled in and modified with the help of SVR regression. (Xiong et al., 2014) developed an offline hybrid model, which is used as a soft sensor to estimate hydrogen consumption rates and fuel gases losses. The hybrid models also assisted for reactor scale up scenarios and kinetic parameters estimation (Jarullah et al., 2011). Hybrid sulfur predictors are utilized effectively for offline monitoring and prediction of sulfur concentration (Kadlec et al., 2009; Bhutani et al., 2006).

In this chapter hybrid model (combination of mathematical and data driven model) will be develop to predict product sulfur concentration for an online HDS system.

5.2 Proposed Hybrid Structure

The proposed hybrid model is shown in Figure 5.2 and it can be explained as follows

In the Figure 5.2, the HDS reactor represents the actual HDS reactor followed by an analyzer (product sulfur analyzer) which analyzes product stream, measures and returns product sulfur concentration. The proposed Hybrid model has two components; (i) an online component (which comprise of data-driven model) and (ii) an offline component (consisting of mathematical model and optimization algorithm).

At any given point of time, process data of feed stream (temperature, feed rate) is preprocessed and fed to the data-driven model along with estimated feed sulfur concentration. The data driven model predicts the product sulfur concentration ($\hat{C}_{Sul}$). The predicted sulfur concentration is updated using previously calculated bias and display for online use. Since the sulfur analyzer is typically located near the tankage there is a delay in obtaining the analyzer measurement. Once the sulfur concentration from the analyzer (C_{Sul}) becomes available, it is compared with predicted sulfur concentration to calculate the error. If the calculated bias changed from the previous stage then there is need to update the predictor. In this present design since feed sulfur is not measured we consider this to be causing the mismatch between the predicted and measured value. When the

difference between (C_{Sul}) and ($\hat{C}_{Sul}$) is beyond the acceptable limit an optimization algorithm is triggered which re-estimates the feed sulfur concentration in an offline mode.

Once optimization algorithm is triggered, online prediction pauses and holds the predicted sulfur concentration to the most recent estimated value. The mathematical model then simulates feed conditions to predict product sulfur concentration ($\hat{C}_{Sul}$). This predicted sulfur concentration ($\hat{C}_{Sul}$) is compared with actual measured product sulfur concentration (C_{sul}). Optimization algorithm then minimizes the difference between (C_{sul}) and ($\hat{C}_{Sul}$) and by changing feed sulfur concentration. This new estimated feed sulfur concentration is passed back to the online component (i.e. data-driven model) along with other feed conditions. Once feed sulfur concentration has been re-estimated the offline component stops and the online prediction process resumes until the next optimization algorithm trigger is activated where the cycle repeats itself.

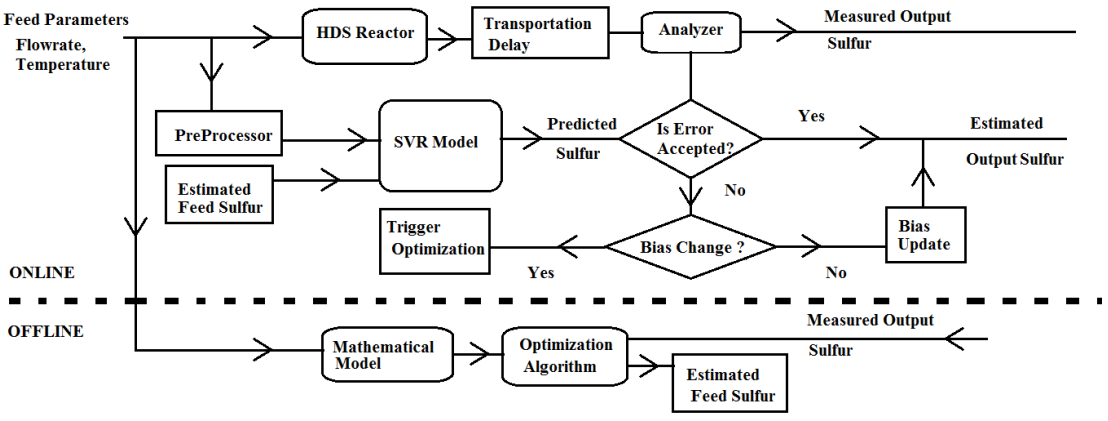

Figure 5.2: Proposed Hybrid Model to Estimate Output Sulfur Concentration for HDS unit

The salient features of the above proposed hybrid model are

- The sulfur predictor is valid for conditions when feed sulfur concentration is varying and not available.
- The proposed hybrid structure is able to predict product sulfur concentration while estimating the feed sulfur concentration in an adaptive manner, without affecting the simulation speed for online prediction.
- Despite being intensive calculation oriented, the mathematical model does not hinder the simulation time of the process, as it is activated only when the optimization algorithm is triggered.

The above hybrid sulfur predictor model comprises of three sub-models

- Mathematical Model (Chapter 3)
- Data Driven Model (Chapter 4)
- Optimization Algorithm

Optimization Algorithm

The Optimization algorithm is an offline component of the proposed hybrid model. When optimization algorithm is triggered, the online simulation pauses and a batch of immediate past historical data set after the change in feed conditions are passed to the mathematical model, which estimates the product sulfur concentration ($\hat{C}$). This algorithm focuses on minimizing the difference between the measured product sulfur

concentration (C) and predicted product sulfur concentration ($\hat{C}$) by mathematical model by calibrating input sulfur concentration.

Mathematically this is represented as the following set of equations

$$J = \min_{C^o} \sum_{k=1}^{q} (C_{Sul,l}^{k,N} - \hat{C}_{Sul,l}^{k,N})^2 \quad \ldots(5.1)$$

where, Sul stands for Sulfur, l is for liquid

$\hat{C}_{Sul,l}^{k,N}$ → estimated product sulfur concentration at the end of the reactor

$C_{Sul,l}^{k,N}$ → measured product sulfur concentration at the end of the reactor

k → number of time points considered in batch window (for our simulation q = 20)

C^0 → feed sulfur concentration.

Subject to the constraints given in Equation (3.1) to (3.8) describing the reactor model. The optimizing parameter feed sulfur concentration enters the mathematical equation through the boundary conditions of the reactor model. After discretization, Equation (3.3) will have the following form at the first discretization point the reactor inlet.

$$\varepsilon_l \frac{C_{Sul,l}^{t+1,1} - C_{sul,l}^{t,1}}{\Delta t} = -u_l \frac{C_{Sul,g}^{t,1} - C_{Sul,g}^{t,0}}{\Delta z} - k_{Sul,g} a_s \left(C_{Sul,l}^{t,0} - C_{Sul,g}^{t,0}\right) \quad \ldots(5.2)$$

and t is for time. The liquid phase sulfur concentration at the boundary $\left(C_{Sul,l}^{t,0}\right)$ is the optimizing parameter, feed sulfur concentration C^0

Also, feed sulfur concentration generally ranges around 5wt% with $\pm$ 2% variation, this forms the initial conditions and limits for our feed sulfur concentration values. The above optimization is carried out using 'fmincon' function from MATLAB optimization toolbox. After estimation of feed sulfur concentration, its value is passed to online SVR data-driven model. The data-driven model resumes predicting product sulfur concentration until a mismatch is detected between the predicted and measured sulfur concentration when the above cycle is repeated.

5.3 Results

In this section, the hybrid model is validated for lab scale and industrial reactor.

5.3.1 Validation for Lab–scale simulated reactor

For validation of lab-scale reactor, we implement hybrid model on the lab –scale reactor, specification of this reactor is mentioned in Table 3.2. The operating conditions are mentioned in Table 4.2 and can be represented in Figure 5.3.

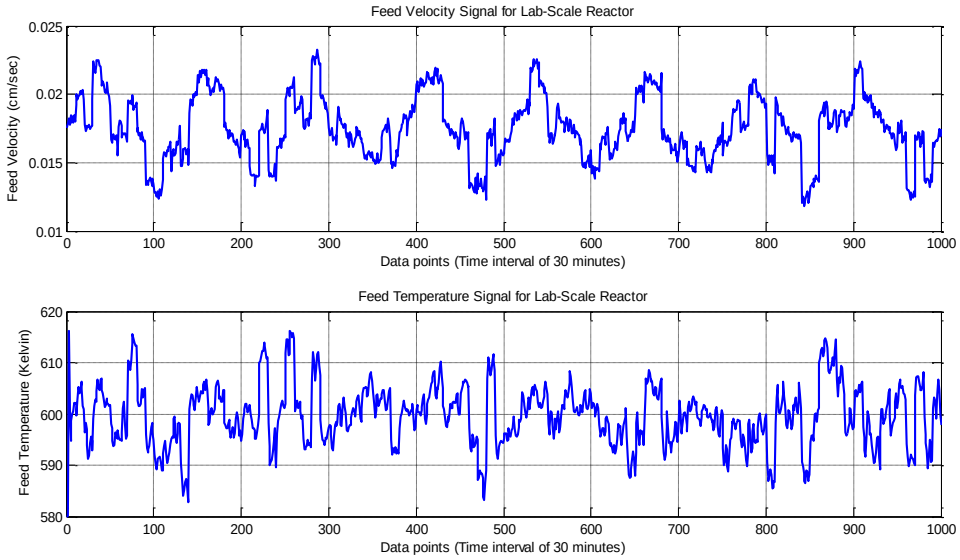

Figure 5.3: Operating conditions signal to simulate Hybrid model for lab-scale reactor

We used these input conditions in mathematical model to generate the historical data set for the measured sulfur concentration. Using the historical data, we implement our hybrid model on our lab-scale reactor. A comparison plot between the measured product sulfur concentration and predicted product sulfur concentration is shown in Figure 5.4.

As the feed sulfur changes, there is a small window, where the model cannot predict the product sulfur. Figure 5.4 shows that transition period where the analyzer data is collected. In the Figure 5.4, there is a small window, where the model cannot predict the product sulfur concentration. This small window is the required time period after which optimization algorithm is triggered, which re-calibrates the feed sulfur concentration. During this window feed sulfur concentration is kept at the most recently estimated value.

At the end of this window, optimization is triggered. Once optimization algorithm is triggered it re-estimates feed sulfur concentration and thus with the new value of feed sulfur concentration, the hybrid model can predict better which is within the acceptable error limit for the model. Our residual plot confirms that the error between the two product sulfur concentration is within acceptable error limits. Thus, our proposed hybrid model can predict product sulfur concentration as close as to measured product sulfur concentration for a lab-scale reactor.

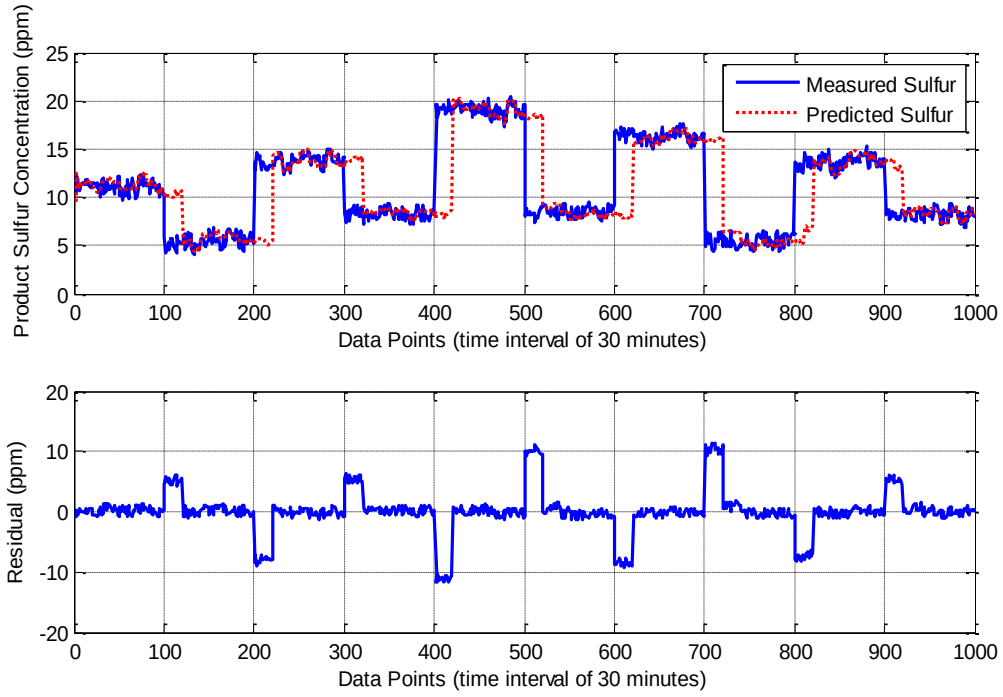

Figure 5.4: Comparison plot Product sulfur concentration for Lab – scale Reactor.

One of the salient feature of proposed hybrid model is that it predict product sulfur concentration while estimating feed sulfur concentration. Therefore, in order to understand the importance estimation of feed sulfur concentration. We implement

proposed hybrid model on the lab-scale reactor with the same historical data, but in this simulation, we do not estimate feed sulfur concentration i.e. optimization algorithm is not activated in this simulation. Figure 5.5 shows the comparison between the measured and predicted product sulfur concentrations when feed sulfur concentration is not estimated.

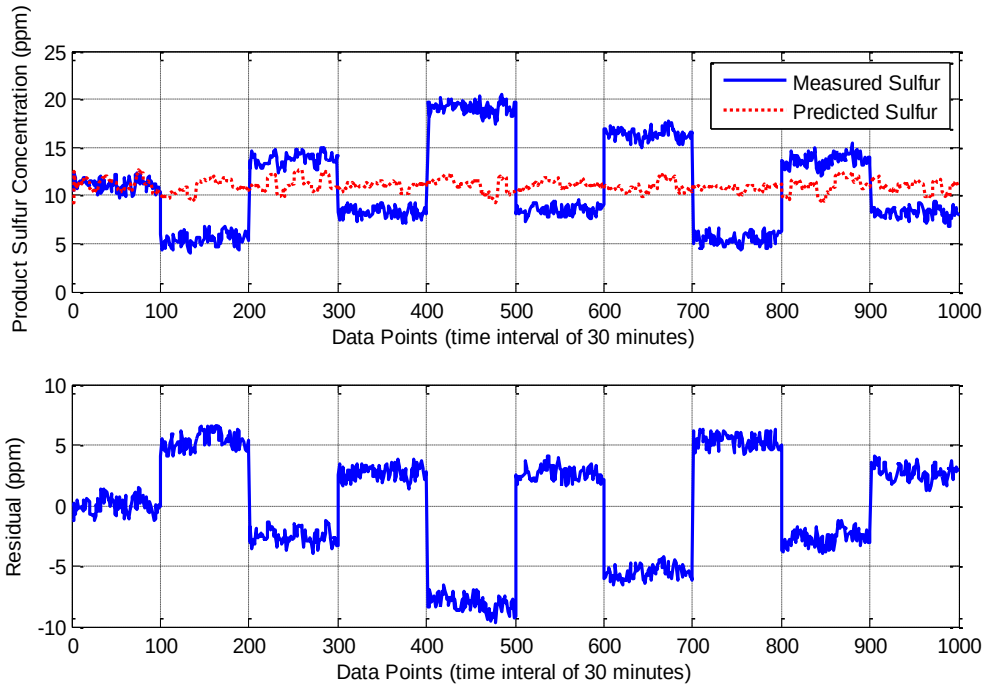

Figure 5.5: Simulation of Hybrid model without estimating Feed Sulfur concentration.

As we can see in Figure 5.5, predicted product sulfur concentration deviates from measured product sulfur concentration. The average error between the two product sulfur concentration is from 10% to 25%. Clearly when feed sulfur concentration is not included in the model prediction is poor. Therefore estimating feed sulfur concentration is essential in order to predict product sulfur concentration effectively.

5.3.2 Validation for industrial plant data.

The specification and feed condition range for the industrial reactor are mentioned in Table 4.3 and Table 4.4 respectively. Figure 5.6 plots the given feed data signal for the industrial reactor. It is evident from given industrial data that feed sulfur concentration is not available. Thus, our historical data consists of velocity of oil, temperature and analyzer sulfur concentration for the industrial reactor. The analyzer is located at the tankage, the measured sulfur will be available with a time delay.

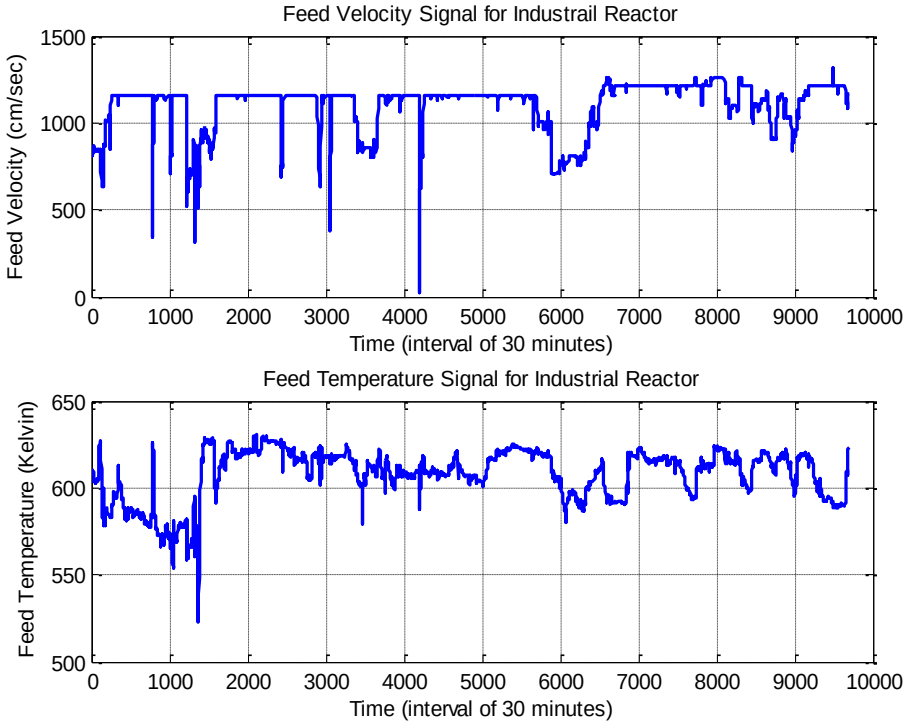

Figure 5.6: Feed Oil Velocity recorded and Feed Temperature recorded for Industrial Reactor

Before we implement our hybrid model for the industrial plant, the industrial data is pre-processed for any flat line and spike so that it will remove necessary data outliers, which might affect the simulation.

We implement proposed hybrid model on the industrial data in order to predict the product sulfur concentration for the HDS unit. A comparison plot of predicted product sulfur and measured product sulfur concentration is shown in Figure 5.7

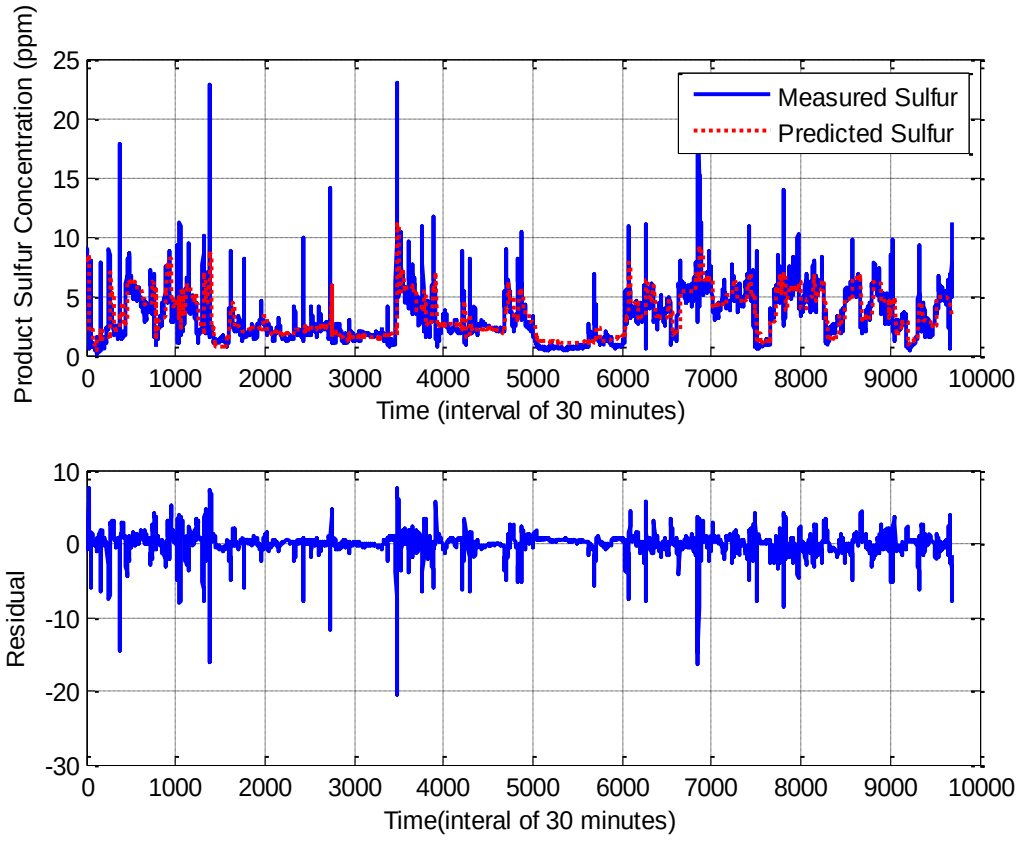

Figure 5.7: Comparison plot for product sulfur concentration for industrial reactor

Figure 5.7 gives a comparison of predicted sulfur concentration and measured sulfur concentration for reactor when feed sulfur concentration values are estimated in the hybrid structure. Our proposed hybrid sulfur reactor was able to predict product sulfur concentration with an average error of 5% to 10% for the entire system. The proposed hybrid sulfur model predictor will be able to predict the product sulfur concentration in an online mode. Thus, it is essential to estimate feed sulfur concentration in order to predict product sulfur concentration as close as to the measured sulfur concentration.

In order to understand significance of optimization algorithm and feed sulfur estimation, in our next simulation we implement the hybrid model but optimization algorithm is not triggered and our hybrid model predicts product sulfur concentration without estimating feed sulfur concentration. Figure 5.8 shows a comparison plot between the predicted and measured product sulfur concentration for the industrial reactor.

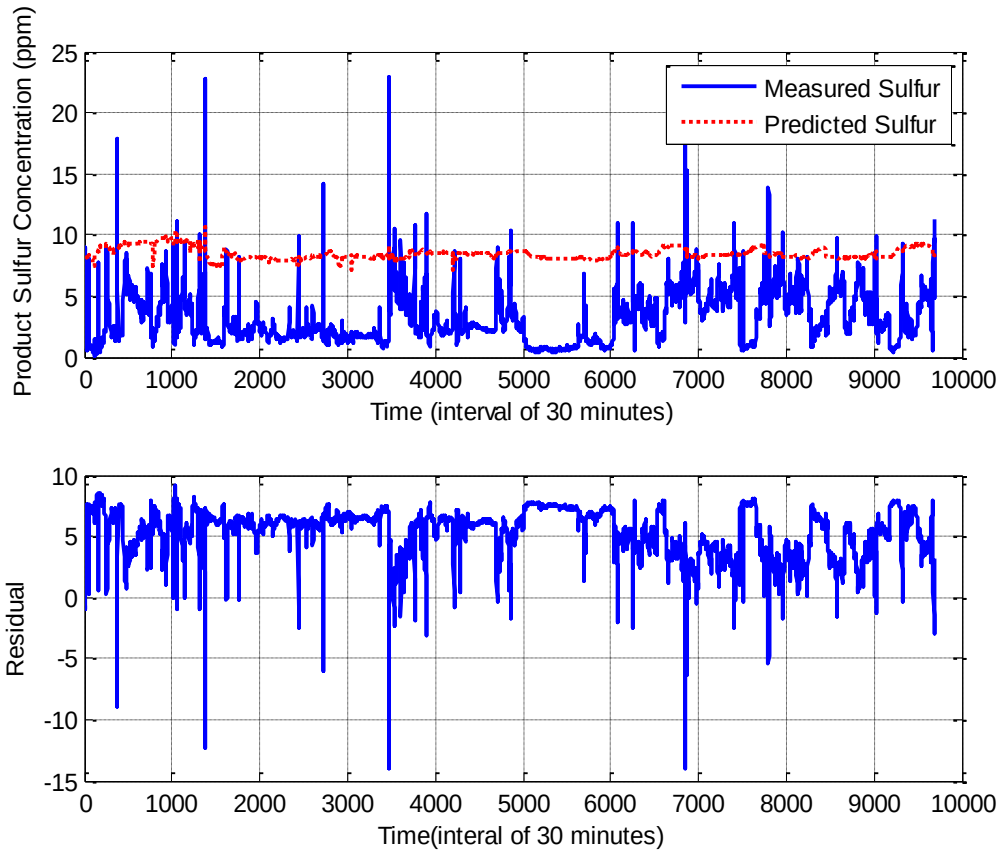

Figure 5.8: Prediction of hybrid model when feed sulfur concentration is not estimated

On comparing Figure 5.7 with Figure 5.8 we could confirm that our hybrid model predictions have higher error when feed sulfur concentration is not estimated (optimization algorithm is not triggered). Thus, it is essential to estimate feed sulfur concentration in order to predict product sulfur concentration as close as to the measured sulfur concentration.

Apart from optimization algorithm, proposed hybrid model perform bias correction if necessary while predicting product sulfur concentration. The bias correction is as

necessary as the optimization algorithm, as it enables our hybrid model to predict product sulfur concentration within error limits.

In this section, we will focus on importance of bias correction. For this simulation, we consider a small section of industrial data where optimization algorithm is not triggered and the hybrid model is predicting product sulfur concentration using online component (surrogate SVR data-driven model). Thus, we consider the 200 data points (data points between 3100 and 3300) from Figure 5.7. The following plot, Figure 5.9 shows the importance of bias correction.

In Figure 5.9, the dotted line is the prediction of the online component. This prediction has some bias error. This error could be due to various reasons like random disturbances, noise or limitation of data-driven model. Hence, this bias error needs to be corrected. Figure 5.9 shows the bias corrected predictions with correction intervals of 6 hours (represented by '-Δ-') and 3 hours (represented by '–x-'). Bias correction improves the prediction of hybrid model as we can see 3-hour biased correction has better prediction results when compared to 6-hour bias correction. Therefore, bias correction assists to improve the prediction of the Hybrid model. As the proposed model is comprises of online component (empirical model), which is developed based on the available historical data and offline component (mathematical model), which has an acceptable fluctuation limit of $\pm$ 20% our model will be able to predict product sulfur concentration effectively even when the input parameters changes. The online component of the proposed hybrid

sulfur predictor is an empirical model. This model is developed using the available historical data. Therefore, the proposed hybrid model is capable to estimate product sulfur concentration within the limits of the historical data.

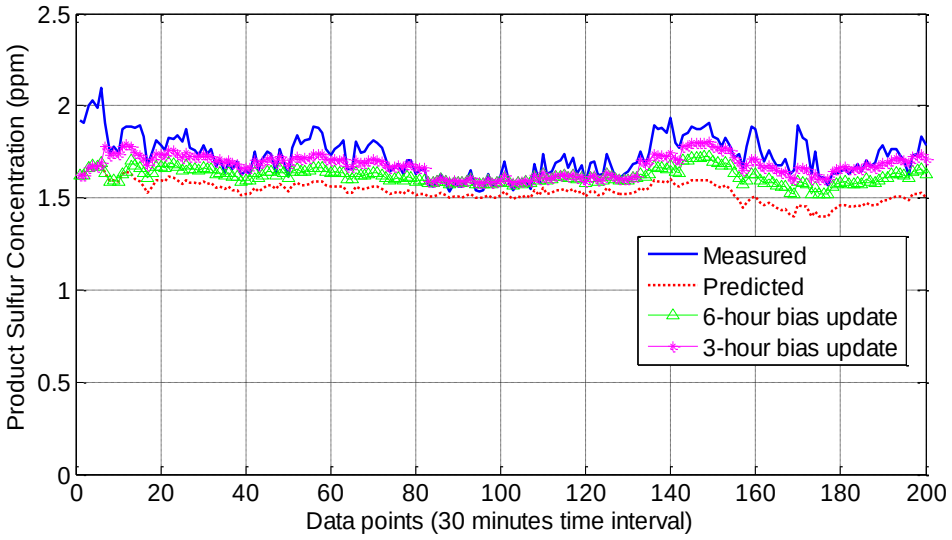

Figure 5.9: Comparison between the measured and predicted sulfur concentration

5.4 Conclusions

In this chapter, we have proposed and explained our hybrid model. The hybrid model proposed has two components (i) online component (which includes surrogate SVR data-driven model explained in chapter 4) and (ii) offline component (includes mathematical model explained in chapter 3 and optimization algorithm). The hybrid model is validated for lab-scale simulated reactor and industrial plant. In both cases, the hybrid model can predict product sulfur concentration within error limits. The significance of optimization algorithm (i.e. estimation of feed sulfur concentration) and bias estimation is also showed in this chapter. The hybrid model tends to have greater error when optimization algorithm

is not triggered and hence optimization algorithm is one of the important components of proposed hybrid model for online product sulfur estimation.

Chapter 6

6 Graphical User Interface

One of the objectives for developing this Hybrid model is to implement it on the online HDS system, in order to predict product sulfur concentration for HDS unit. In this chapter, a Graphical User Interface (GUI) is developed which would assist our proposed hybrid model to be implemented on any given HDS system for online sulfur prediction. MATLAB is used to develop the GUI and it is stated and explained below.

6.1 Developed Graphical User Interface for Online Estimation

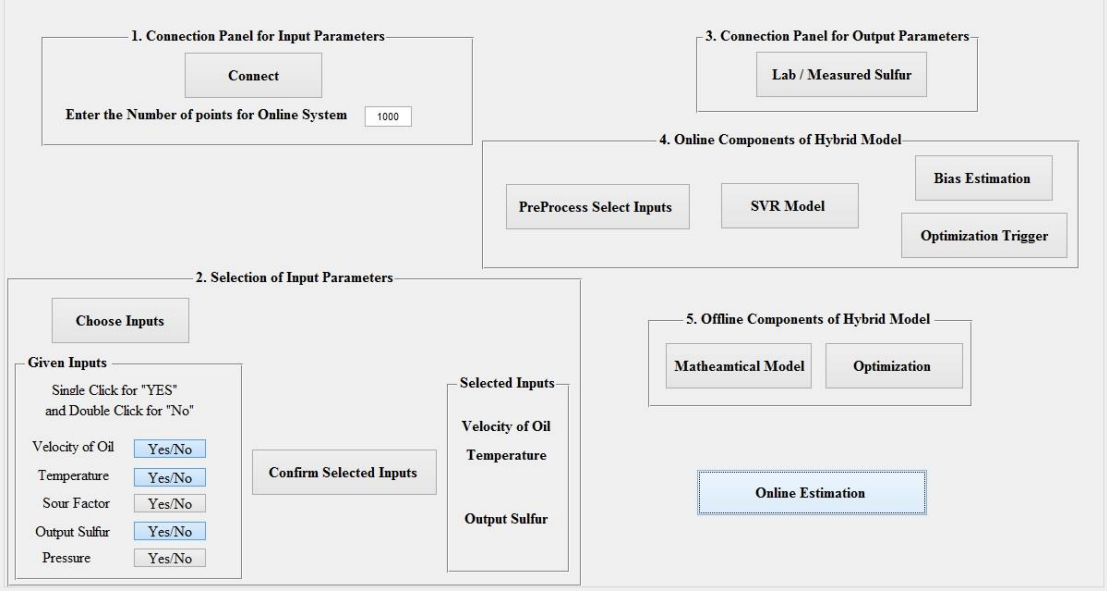

Figure 6.1: Graphical User Interface designed for Proposed Hybrid Model Implementation for Online System

The Figure 6.1 displays the Graphical User Interface (GUI), which is designed for proposed Hybrid model (mentioned in chapter 5) to estimate online product sulfur concentration of fuel. The major components of this GUI is stated as follows

- Connection Panel for Input Parameters
- Selection of Input Parameters
- Connection Panel for Output Parameters
- Online Components of Hybrid Model
- Offline Components of Hybrid Model
- Online Estimation

6.1.1 Connection Panel for Input Parameters

The primary goal of this tab is to connect the hybrid model with the online HDS data stored in the system. The “Connect” Tab looks for the excel file on the system in which online HDS system is stored. Once activated/clicked on “Connect” Tab it requests the user to input the number of data-points, one wish to simulate for sulfur prediction. This is necessary, as HDS process is a continuous process; therefore, the data is continuously recorded in the system and hence for sulfur predictor, there should be a defined endpoint. Once the user key in the number of points, it marks as the endpoint for our hybrid model simulation. Once this endpoint is reached the GUI stops working and final results are displayed. After which proposed GUI can be implemented again for a new section of online data simulation. This helps in online simulation as the user will have complete control as in when to implement sulfur predictor for online product sulfur estimation.

6.1.2 Selection of Input Parameters

In an industrial scenario, various parameters are recorded and stored which may or may not be required by the hybrid model in order to predict the sulfur concentration. Therefore, GUI gives an option to the user to select the necessary parameters, which will assist in simulating the hybrid model. When clicked on “Choose Inputs” tab “Given Inputs” panel is displayed for the user which will list all the parameters which are stored in the online data file for HDS process. The user then choose, necessary parameters by single clicking on the “Yes/No” tab which is beside the parameter name. As mentioned in the panel Single click on “Yes/No” tab consider the parameter for the simulation while double clicking on “Yes/No” tab ensure that the parameter is not selected for simulation. When satisfied with the selection of the parameters the user then confirm their choice by clicking on “Confirm Selected Inputs” tab, which will display another panel named “Selected Inputs” consisting all the parameters, which are selected by the user. In Figure 6.1, the parameters, which are selected for hybrid model simulation, are Velocity of Oil, Temperature (Inlet) and Product sulfur Concentration and thus they are displayed in “Selected Inputs” parameters. The selected parameters are utilized by the GUI interface to predict online product sulfur concentration.

6.1.3 Connection Panel for Output Parameters

By now, GUI has created a temporary file where the selected parameters are stored and the endpoint for the simulation is marked. When clicked on “Lab/Measured Sulfur” tab, GUI selects product sulfur concentration values recorded by the analyzer and it is stored

in as the separate variable for GUI. As we can see in the above figure, the product sulfur concentration values are stored by GUI in two different location, one of them is stored in the “Selection of Input Parameters” Panel and other is stored by “Connection Panel for Output Parameters” panel. The values stored by both these panels are same. “Connection Panel for Output Parameters” is just a failsafe switch, which will ensure that if the user missed to include product sulfur concentration as one of the parameters for simulation, this panel will record the value for GUI in order to simulate the hybrid model.

6.1.4 Online Components of Hybrid model.

As mentioned in our hybrid model (Chapter 5), the online component comprises of our Support Vector Regression Model (SVR) and Bias Calculation. These elements have different tab in GUI in order to work efficiently and are stated as follows

- “Preprocess Select Inputs” Tab
- “SVR Model” Tab
- “Bias Estimation” Tab
- “Optimization Trigger” Tab

Preprocess Select Inputs Tab

In industrial process, preprocessing is one of the important step, for simulation. This is because after the start-up phase of the plant, it operates in more or less constant and fully optimal state. However, even the steady state is progressively changing with time. The

most common causes of the process operating point changes are the changes of the process product demand, the change of the catalyst activity, clogging of heat exchangers, etc. Due to which the data recorded tends to have some disturbance or irregularities, which needs to be eliminated before simulation. As this data with noise will cause anomalies, which will either hinder or abruptly stop the simulation process (Kadlec et al., 2009). While preprocessing, GUI looks out for two types of error and rectifies them

- Flat Line
- Spike

Flat Line

Flat Line Detection is a classic case of Missing data scenario. Missing data are single samples or consequent sets of samples, where one or more variables (i.e. measurements) have a value, which does not reflect the real state of the physical measured quantity. The affected variables usually have values like $\pm\infty$, 0 or any other constant value. Possible causes of missing data are related to the transmission of the data between the sensors and the database, errors in the database, problems in accessing the database, etc. There are various approaches, which could eliminate flat line once detected. A non-optimal approach is to skip the data samples consisting of variable or variables with the missing values, i.e. case deletion. Another approach to missing values handling takes into account the multivariate statistics of the data and thus makes the reconstruction of the missing values dependent on the other available variables of the affected samples. However, most

commonly used approach is, to replace the missing values with the mean values of the affected variable, which is used by GUI here.

Spike

Spike Detection is another form of Data Outliers. Outliers are sensor values, which deviate from the typical or sometimes also meaningful, ranges of the measured values. These are the values, which violate the physical or technological limitations. For example, the absolute pressure may not reach negative values or flow sensor may not deliver values, which exceed the technological limitations of the sensor. To be able to detect this type of outlier efficiently the system has to provide with the limiting values in the form of a priori information. In order to eliminate these spikes, the most commonly used approach is, to replace the spiked values with the mean values of the affected variable.

Once the “PreProcess Select Inputs” tab is clicked the GUI go through the data and looks for flat line and spike once detected it merely replace the value with the mean. This ensures that while simulating our regression model does not comes across values like $\pm \infty$ or 0 or any numerical constant which lies outside the range of the process data. This preprocessed data passes to the Online SVR regression tab for prediction of output sulfur concentration.

SVR Model tab

When clicked on this tab, preprocessed data and feed sulfur concentration is passed to surrogate SVR data driven model (which is developed and validated in Chapter 4). This is where the surrogate SVR data driven model, predicts the product sulfur concentration for the given feed data. After which the data is passed to "Bias Estimation" Tab.

Bias Estimation Tab

The bias estimation tab compares the product sulfur concentration estimated by SVR model for online prediction with actual product sulfur concentration data (which is measured by sulfur analyzer), to calculate error. If the error is more than acceptable limit, a signal is sent to optimization Trigger tab. If not bias update is performed (if necessary) on SVR predicted online product concentration which is then displayed as estimated product sulfur concentration for HDS unit.

Optimization Trigger Tab

Once the tab is activated / clicked, it accepts the signals, which is passed from the "Bias Estimation" Tab and upon receiving the input. Optimization Trigger Tab pauses the online simulation and activates the mathematical model and optimization tabs in offline mode.

6.1.5 Offline Components of Hybrid Model

As mentioned in chapter 5, the offline part of proposed Hybrid model consists of mathematical model and optimization algorithm. The sole objective of this offline model is to re-estimate feed sulfur concentration, which could be passed to the online part along with the selected input parameters to predict appropriate output sulfur concentration. Thus, this Offline components panel has two tab namely Mathematical Model Tab and Optimization tab.

Mathematical Tab

This tab gets activated after Optimization Trigger is triggered in the Online Components of Hybrid model. Once the online prediction pauses, current preprocessed feed conditions are simulated by mathematical model, estimating product sulfur concentration, which is passed to the optimization Tab

Optimization Tab

This tab optimizes the difference between the two product sulfur concentration (measured and predicted). This optimization is subjected to HDS unit equation and chemical kinetics to re – estimate feed sulfur concentration. Once feed sulfur concentration is re-estimated, optimization tab passes these new value to surrogate SVR model and stops.

6.1.6 Online Estimation Tab.

Online estimation ensures that the cycle keeps on repeating until user-defined endpoint is passed. Once the endpoint is reached, it plots a graph between the estimated online product sulfur concentration and measured product sulfur concentration. Based on this graph the controller can take the necessary action, which would ensure that the fuel has a lower sulfur concentration value as per the guidelines for the refining process.

6.2 Results

In this section GUI, simulates the Hybrid model for the give set of industrial data. The industrial data for which the hybrid model will be implemented is a collective data set for a period of more than a year. The screen shot of GUI, which acts as the interface, to implement our proposed hybrid model is represented in Figure 6.1.

As, GUI completes the simulation by Hybrid model and estimating the output sulfur concentration for the entire data series, Figure 6.2 represents the comparison plot between the predicted product sulfur concentration and measured product sulfur concentration .

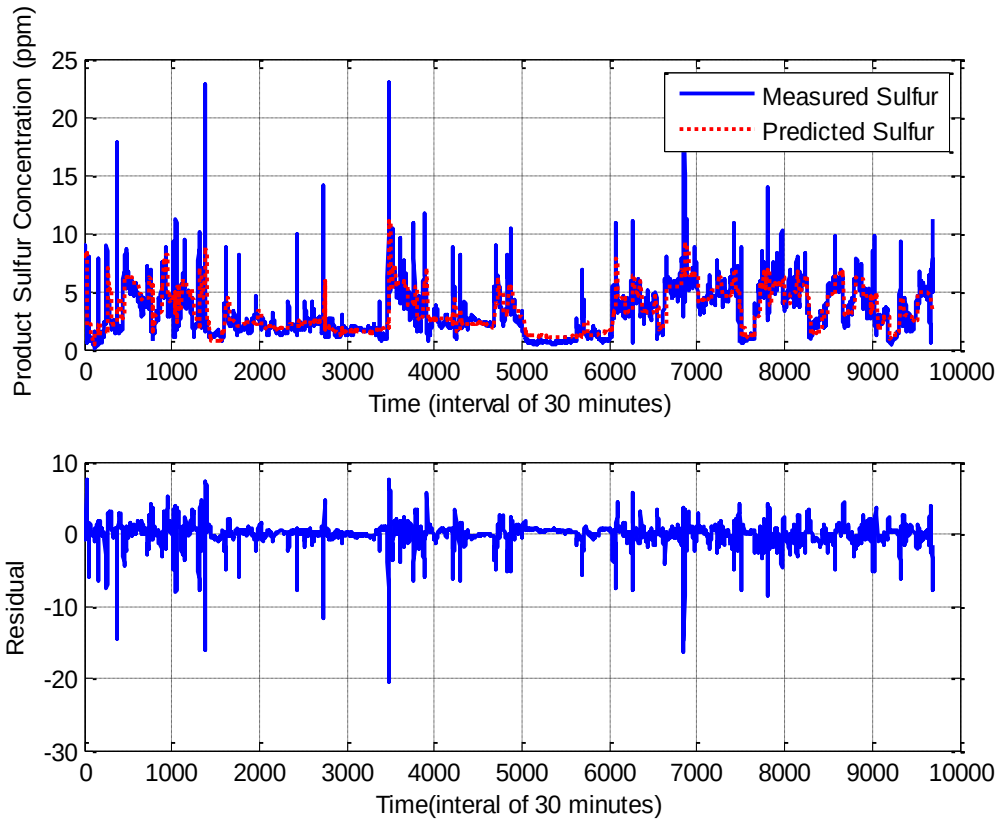

Figure 6.2: Plot generated by GUI comparing Product Sulfur Concentration.

From Figure 6.2 we could conclude that our GUI is able to simulate the hybrid model satisfactorily and the plots generated by the GUI concludes that proposed hybrid model can estimate output sulfur concentration with 10% of error limit.

6.3 Conclusions

In this chapter, we had developed a Graphical User Interface (GUI), which can be implemented on any HDS system. The proposed GUI when simulated, implements the hybrid model to estimate product sulfur concentration for online system. GUI while

implementing hybrid model, removes time lag, present in measured product sulfur concentration by sulfur analyzer. It also pre-process industrial data before the data is being used by hybrid model for product sulfur prediction. Once the marked end of GUI is achieved, it plots a graph comparing predicted product sulfur concentration with measured product sulfur concentration. GUI generates a plot comparing the estimated product sulfur concentration and measured product sulfur concentration.

Chapter 7

7 Conclusion

In this thesis, a hybrid model for hydrodesulfurization is proposed and validated using lab-scale reactor and industrial data. The salient features of the proposed hybrid model are; (i) the sulfur predictor is valid for conditions when feed sulfur concentration is varying and not available. (ii) The proposed hybrid structure is able to predict product sulfur concentration while estimating the feed sulfur concentration in an adaptive manner, without affecting the simulation speed for online prediction and (iii) Despite being intensive calculation oriented, the mathematical model does not hinder the simulation time of the process, as it is activated only when the optimization algorithm is triggered. The proposed hybrid structure is a combination of mathematical model, optimization algorithm (off-line mode) and SVR data-driven model (on-line mode). The hybrid model can predict online product sulfur concentration with an error of 5% to 10%, thus giving a better opportunity to control the desulfurization process. The hybrid sulfur predictor is configured to predict product sulfur concentration as close as to measured product sulfur concentration while estimating feed sulfur concentration, which is unknown for the desulfurization unit. The SVR model predictions are improved with help of CCD design – a DOE technique to ensure that our online product sulfur predictions are as close as to the measured product sulfur concentration. The salient feature of this hybrid model is that it can estimate feed sulfur concentration adaptively which keeps on fluctuating for a desulfurization unit. In addition, as the mathematical model is in offline mode it does not

affect the online simulation with its processing time, unlike previous models in the literature. The MATLAB based GUI developed for sulfur predictor makes it easy to deploy on any desulfurization unit for prediction purpose.

7.1 Future Scope of Work

The future studies, some of the limitations of this model can be eliminated. The biggest limitation of the proposed model is similar to the empirical models developed in past. As the online component of proposed sulfur predictor is developed using available historical data, the data selected to develop the empirical model should not have any irregularities or noise. This will ensure that the empirical model is able to predict product sulfur concentration with a higher accuracy. The recorded time interval between two data points in the industrial data if reduced will assist to develop an effective and more sensitive hybrid sulfur predictor. This smaller time interval between two data points will reduce the acceptable window of error which triggers optimization algorithm and thus our predictor will be able to predict product sulfur concentration more faster and effectively. In the offline component where numerous equations are used to calculate various parameters, if we actually measured these parameters directly from the field, process it will reduce the simulation time for offline component and the results will have more accuracy. The proposed hybrid model should be validated with other industrial data before we implement to actual HDS unit, this will validate the prediction capabilities of our proposed hybrid sulfur predictor.

Reference

Ahmed Faisal, Nazir Salman, and Yeong Koo Yeo, *A recursive PLS-based soft sensor for prediction of the melt index during grade change operations in HDPE plant*, Korean Journal of Chemical Engineering, **2009**, vol.26, no.1, p.14.

Alvarez, Anton, Ancheyta J., and Muñoz J.A.D., *Modeling, simulation and analysis of heavy oil hydroprocessing in fixed-bed reactors employing liquid quench stream*, Applied Catalysis A, **2009**, vol.361, no.1, p.1-12.

Ancheyta, J. Mederos, F.S., Rodrı´guez, M.A. and Arce, E., *Dynamic Modelling and Simulation of Catalytic Hydrotreating Reactors*, Energy & Fuels, **2006**, vol.20, p.936.

Avraam Dimitrios G., and Vasalos I.A., *HdPro: a mathematical model of trickle-bed reactors for the catalytic hydroprocessing of oil feedstock*, Catalysis Today,**2003,** vol.79, p.275-283.

Babich I. V., and Moulijn J.A., *Science and technology of novel processes for deep desulfurization of oil refinery streams: a review*, Fuel, **2003**, vol.82, no.6, p.607-631.

Bahzad, Dawoud, and Ezra Kam, *Atmospheric residues of Kuwait heavy crude upgrading: a pilot plant and model simulation study*, Energy & Fuels, **2009**, vol.23, no.3, p.1683-1694.

Behnasr M. and Jazayeri-Rad H., *Robust data-driven soft sensor based on iteratively weighted least squares support vector regression optimized by the cuckoo optimization algorithm*, Journal of Natural Gas Science and Engineering, **2015**, vol. 22, p. 35–41.

Berger D., Landau M.V., Herskowitz M., Boger Z., *Deep Hydrodesulfurization of Atmospheric Gas Oil—Effects of Operating Conditions and Modelling by Artificial Neural Network Techniques*, Fuel, **2006**, vol.75, no.7, p.907–911.

Bhaskar, M., Valavarasu G., Sairam B, Balaraman K.S. and Balu K., *Three-phase reactor model to simulate the performance of pilot-plant and industrial trickle-bed reactors sustaining hydrotreating reactions*, Industrial & engineering chemistry research, **2004**, vol.43, no.21, p.6654-6669.

Bhutani N., Rangaiah G. P., and Ray A. K., First-Principles , Data-Based , and Hybrid Modeling and Optimization of an Industrial Hydrocracking Unit, Industrial Engineering Chemistry Research, **2006**, vol.45, p. 7807.

Bro, R. Multiway calibration. Multilinear PLS Journal of Chemometrics, **1996**, vol.10, no.1, p.47–61.

Box G.E.P., and Wilson K.B., *On the Experimental Attainment of Optimum Conditions (with discussion)*, Journal of the Royal Statistical Society Series B, **1951**, vol.13, no.1, p.1.

Cicili P., Reddy G. P. and Kumar V. R. *Dynamic matrix control for HDS reactor*. In Proc. of the International Multi-conference of Engineers and Computer Scientists,**2009**, Hong-Kong.

Chen, Jinwen, Vinay Mulgundmath, and Neil Wang, *Accounting for vapor− liquid equilibrium in the modeling and simulation of a commercial hydrotreating reactor*, Industrial & Engineering Chemistry Research, **2010**, vol.50, no.3, p.1571-1579.

Cheng, Z.-M., Fang, X.-C., Zeng, R.-H., Han, B.-P., Huang, L., Yuan, W.-K., *Deep Removal of Sulfur and Aromatics from Diesel Through Two-Stage Concurrently and Countercurrently Operated Fixed-Bed Reactors*, Chemical Engineering Science,**2004**, vol.59, no.22, p.5465–5472.

Chowdhury, R., Pedernera, E. and Reimert, R., *Trickle-Bed Reactor Model for Desulfurization and Dearomatization of Diesel*, American Institute of Chemical Engineers, **2002**, vol.48, no.1, p 126

Dayal, B. S., & MacGregor, J. F., *Recursive exponentially weighted PLS and its applications to adaptive control and prediction*, Journal of Process Control, **1997**, vol.7, no.3, p.169–179.

De la Paz-Zavala C., Burgos-Vázquez E., Rodríguez-Rodríguez J.E., and Ramírez-Verduzco, *Ultra low sulfur diesel simulation. Application to commercial units*, Fuel, **2013**, vol. 110, p.227

Deng Z., Wang T., and Wang Z., *Hydrodesulfurization of diesel in a slurry reactor*, Chemical Engineering Science, **2010**, vol.65, no.1, p. 480–486.

Dong, Yaming, and Xuefeng Yan, *A comparative study of hybrid models combining various kinetic and regression models for p-xylene oxidation*, Korean Journal of Chemical Engineering, **2014**, vol.31, no.10, p.1746.

Dudukovic M.P.,Larachi F., Mills P.L., *Multiple catalytic reactors: a perspective on current knowledge and future trends*, Catalysis Review, **2002**, vol.44, no.1, p.123‑246

Estrada-Flores I. Merts B. De Ketelaere J., Lammertyn, *Development and validation of "grey-box" models for refrigeration applications: a review of key concepts*, Int. J. Refrig., **2006**, vol. 29, p.931.

EU Fuel Regulations: http://www.dieselnet.com/standards/eu/fuel.php.

Feng, Rui, Shen, Wei, Shao, Huihe, *A soft sensor modeling approach using support vector machines,* Proceedings of the. In: American Control Conference, IEEE, Atlanta, June 1-3, **2003**.

Fujikawa, Takashi, *Highly active HDS catalyst for producing ultra-low sulfur diesel fuels*, Topics in Catalysis, **2009**, vol. 52, no.6-7, p.872.

Farahani, Hadi F., and Shahrokh Shahhosseini, *Simulation of Hydrodesulfurization Trickle Bed Reactor*, Chemical Product and Process Modeling, **2011**, vol.6. no.1.

Gernaey K.V. and Gani R., *A model-based systems approach to pharmaceutical product-process design and analysis*, Chemical Engineering Science, **2010**, vol. 65, no. 21, pp. 5757–5769.

Gonzalez G.D., *Soft sensors for processing plants*, Intelligent Processing and Manufacturing of Materials, IEEE, **1999**, vol. 44, no. 3, p. 893.

Gonzaga, J. C. B., Meleiro L.A.C., Kiang C. and Maciel Filho R., *ANN-based soft-sensor for real-time process monitoring and control of an industrial polymerization process*, Computers and Chemical Engineering, **2009**, vol.33, no.1 p.43-49.

Henrik, Topsøe, Clausen Bjerne S., and Massoth F.E., *Hydrotreating catalysis*, Catalysis, **1996**, p. 1- 269, Springer Berlin Heidelberg.

Jones D. S. and Pujadó P. P. (Eds.), *Handbook of petroleum processing*, Springer Science & Business Media, **2006**.

Japan Fuel Regulations: http://www.dieselnet.com/standards/jp/fuel.php.

Jarullah, Aysar T., Iqbal M. Mujtaba, and Alastair S. Wood, *Kinetic parameter estimation and simulation of trickle-bed reactor for hydrodesulfurization of crude oil*, Chemical Engineering Science, **2011**, vol.66, no.5, p.859.

Jolliffe I. , *Principal component analysis*, John Wiley and Sons, **2005**, p.225.

Kadlec P., Gabrys B., and Strandt S., *Data-driven Soft Sensors in the process industry*, Computer and Chemical Engineering, **2009**, vol.33, no.4, p.795–814.

Khalfhallah H. A., *Modelling and optimization of oxidative desulfurization process for model sulfur compounds and heavy gas oil*, Diss. PhD Thesis, University of Bradford, **2009**.

Kim T., Ali S. A., Alhooshani K., Park J.I., Al-Yami M.,Yoon S.H., and Mochida I., *Analysis and deep hydrodesulfurization reactivity of Saudi Arabian gas oils*, Journal of Industrial and Chemical Engineering, **2013**, vol.19, no.5, p.1577–1582.

Knudsen, Kim G., Barry H. Cooper, and Henrik Topsøe, *Catalyst and process technologies for ultra low sulfur diesel,* Applied Catalysis , **1999**, vol.189, no.2, p.205-215.

Kourti T, *Process analysis and abnormal situation detection: From theory to practice. Control Systems Magazine,* IEEE, **2002**, vol.22, no.5, p.10–25.

Korsten, H. and Hoffmann, U., *Three-Phase Reactor Model Pilot Trickle-Bed for Hydrotreating in Reactors*, American Institute of Chemical Engineeirng, **1996**, vol.42, p.1350.

Lukec K. Sertić-Bionda, and Lukec D., *Prediction of sulphur content in the industrial hydrotreatment process*, Fuel Processing Technology, **2008**, vol. 89, no. 3, p.292–300.

Laird T., *Book review of pharmaceutical dosage forms: tablets*. Org Process Res Dev, **2009**, vol.13, no.3, p.656.

Lye L., *ENGI 9516: Similitude, modeling and data analysis – course notes*, Memorial University of Newfoundland and Labrador, **2013**, St.John's,Canada

Lou H.C., Su H.Y., Xie L., Gu Y., Rong G., Inferential model for industrial polypropylene melt index prediction with embedded priori knowledge and delay estimation, Industrial & Engineering Chemistry Research, **2012**, vol.51, no.25, p.8510.

Marafi, A., Al-Hindi A., and Stanislaus A., *Deep desulfurization of full range and low boiling diesel streams from Kuwait Lower Fars heavy crude*, Fuel Processing Technology, **2007**, vol.88, no.9, p.905.

Mederos F.S. and Ancheyta J., *Mathematical modeling and simulation of hydrotreating reactors: Cocurrent versus countercurrent operations*, Applied Catalyst, **2007**, vol. 332, no.1, p.8.

Mederos F.S., Elizalde I., and Ancheyta J., *Steady-State and Dynamic Reactor Models for Hydrotreatment of Oil Fractions: A Review*, Catalyst Review, **2009**, vol. 51, no. 4, p.485–607.

Macı´as M.J. and Ancheyta J., *Simulation of an isothermal hydrodesulfurization small reactor with different catalyst particle shapes*, Catalyst Today, **2004**, vol.98, p.243.

Montgomery D. C. Johnson R. T., Jones B. and Parker P. A., *Comparing computer experiments for fitting high-order polynomial metamodels*, Journal of Quality Technology, **2010**, vol.42, p.86-102.

Montgomery D.C., *Design and Analysis of Experiments*, **2010**, 7th Edition. John Wiley and Sons, New York, United States of America.

Mochida, Isao, *Deep hydrodesulfurization of diesel fuel: Design of reaction process and catalysts*, Catalysis Today, **1996**, vol.29, no.1, p.185.

Mat Noor R.A., Ahmad Z., Mat Don M., Uzir M.H., *Modelling and control of different types of polymerization processes using neural networks technique: a review.* Canadian Journal of Chemical Engineering, **2010**, vol.88, no.6, p.1065–1084.

Ocic, Ozren, Oil refineries in the 21st century: energy efficient, cost effective, environmentally benign. John Wiley & Sons, **2010**, p.289

Oyama, Ted S., *Transition metal phosphide hydroprocessing catalysts: a review*, Catalysis Today, **2009**, vol.143, no.1, p.94.

Principle, Jose C., Euliano, Neil R., Curt Lefebvre, *Neural and Adaptive Systems*, John Wiley & Sons, **2000**, p.529

Porto S., Novel Strategies for Process Control Based on Hybrid Semi-parametric Mathematical Systems, PdD Thesis, Universidade do Porto, Porto, **2011.** p.10.

Qin S. J., *Recursive PLS algorithms for adaptive data modeling*, Computers and Chemical Engineering, **1998**, vol.22, no.4–5, p.503.

Rodriguez M.A., and Ancheyta J., Modeling of hydrodesulfurization (HDS), hydrodenitrogenation (HDN), and the hydrogenation of aromatics (HDA) in a vacuum gas oil hydrotreater, Energy & fuels, **2004**, vol.18, no.3, p.789.

Sliskovic, Drazen, Grbic, Ratko, Hocenski, Zeljko, *Methods for plant data-based process modeling in soft-sensor development*, IEEE, **2011**, vol.20, no.11, p.2097.

Shi, J. and Liu X.G., *Melt index prediction by neural soft-sensor based on multiscale analysis and principal component analysis*, Chinese Journal of Chemical Engineering, **2005**, vol.13, no.6, p.849–852.

Sharmin R., Sundararaj U., Shah S., Griend L.V., Sun Y.J., *Inferential sensors for estimation of polymer quality parameters: Industrial application of a PLS-based soft sensor for a LDPE plant*, Chemical Engineering Science, **2006**, vol.61, no.19 p.6372-6384.

Shokri S., and Zarrinpashne S., A Mathematical model for calculation of effectiveness factor in catalyst pellets of hydrotreating process, Petroleum & Coal, **2006**, vol.48, p.27.

Slimani, Mohammed Zakaria, Dynamic heterogeneity in random and gradient copolymers: A computational investigation, Macromolecules, **2013**, vol.46., no.12, p. 5066.

Speight, G.James, *Handbook of petroleum product analysis*, John Wiley & Sons, **2015,** p 189.

Stanislaus A., Al-Barood A., and Qabazard H., *Effect of Feed Quality on Deep Desulfurization of Diesel*, Preprints of Paper American Chemical Society Divison of Fuel Chemistry, **2003**, vol.48, no.2, p. 655.

Stratiev D, Galkin V., and Stanulov K., *Study: Most-active catalyst improve's ULSD economics*, Oil & gas journal, **2006**, vol.104, no.30, p.53-57.

Thompson M.L., Kramer M.A., *Modeling chemical processes using prior knowledge and neural networks*, Am. Inst. Chem. Eng. J., **1994**, vol.40, p.1328.

Vergov, I., and I. Shishkova, *Catalyst advances promote production of near zero sulphur diesel*, Petroleum & Coal, **2009**, vol.51, no.2, p.136-139.

Vapnik, Vladimir, *The Nature of Statistical Learning Theory*, Springer, **2000**, Verlag, NewYork.

Von Stosch M., Oliveira R. , Peres, S. Feyo J. , and Azevedo D., *Hybrid semi-parametric modeling in process systems engineering : Past , present and future,* **2014**, vol. 60, p. 86.

Wold, Svante, Sjostroma, Michael, Erikssonb, Lennart, *PLS-regression: a basic tool of chemometrics*, Chemom. Intelligent Laboratory Syst., **2001**, vol.58, p.109.

Xiong, Weili, Zhang, Wei, Liu, Dengfeng, Xu, Baoguo, *Fuzzy pruning based LS-SVM modeling development for a fermentation process*, Applied Analysis, **2014**, vol. 1, p.7.

Yan, Weiwu, Shao, Huihe, Wang, Xiaofan, *Soft sensing modeling based on support vector machine and Bayesian model selection*, Computer Chemical Engineering, **2004**, vol.28, no.8, p.1489.

Yin, Shen, Wang, Guang, Karimi, Hamid Reza, *Data-driven design of robust fault detection system for wind turbines*, Mechatronics, **2014**, vol.24, no.4, p.298 - 306.

Zhang, J., Jin, Q., Xu, Y.M., *Inferential estimation of polymer melt index using sequentially trained bootstrap aggregated neural networks,* Chemical Engineering & Technology, **2006**, vol.29, no.4, p. 442–448.